

# Pricing and Risk Management in Competitive Electricity Markets

A Thesis  
Presented to  
The Academic Faculty

by

**Zhendong Xia**

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

School of Industrial and Systems Engineering  
Georgia Institute of Technology  
December 2005

# Pricing and Risk Management in Competitive Electricity Markets

Approved by:

Dr. Shijie Deng, Advisor  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Dr. Paul Griffin  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Dr. David Goldsman  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Dr. A.P. Sakis Meliopoulos  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Dongjun Wu  
College of Management  
*Georgia Institute of Technology*

Date Approved: November 17, 2005

*Dedicated to,*  
*my parents and Angie Zhang,*

## ACKNOWLEDGEMENTS

My first, and most earnest, acknowledgment must go to Dr. Shijie Deng, my principal advisor for his enthusiastic and continuous support and mentorship. The second person I like to thank is Dr. Liang Peng who provides me a lot of his valuable suggestions and helps on my work. I would also like to thank the rest of my thesis committee: Dr. David Goldsman, Dr. Paul Griffin, Dr. Dongjun Wu and Dr. Athan Meliopoulos who took effort in reading and providing me with insightful comments and precious advice. I would like to thank all my friends for their continued moral support.

Finally, I am forever indebted to my wonderful parents and Angie Zhang for their understanding, endless patience and encouragement when whenever I needed. They deserve far more credits than I can ever give them.

# TABLE OF CONTENTS

<b>DEDICATION . . . . .</b>	<b>iii</b>
<b>ACKNOWLEDGEMENTS . . . . .</b>	<b>iv</b>
<b>LIST OF TABLES . . . . .</b>	<b>vii</b>
<b>LIST OF FIGURES . . . . .</b>	<b>viii</b>
<b>SUMMARY . . . . .</b>	<b>x</b>
<b>I OVERVIEW . . . . .</b>	<b>1</b>
1.1 Electricity Spot Prices Modeling . . . . .	3
1.2 Market-Based Valuation . . . . .	5
1.3 GARCH Models and Value at Risk . . . . .	8
1.4 Contributions of this Work . . . . .	10
<b>II ELECTRICITY SPOT PRICES MODELLING . . . . .</b>	<b>12</b>
2.1 Introduction . . . . .	12
2.2 Variance Gamma Process Review . . . . .	14
2.3 Mean-Reverting Variance Gamma Process . . . . .	16
2.4 Simulation Study . . . . .	17
2.4.1 Model Parameters Estimation . . . . .	17
2.4.2 Derivatives Pricing . . . . .	22
2.5 Conclusion . . . . .	26
<b>III PRICING AND HEDGING ELECTRICITY SUPPLY CONTRACTS: A CASE WITH TOLLING AGREEMENTS . . . . .</b>	<b>29</b>
3.1 Introduction . . . . .	29
3.2 Problem Description . . . . .	30
3.3 A Stochastic Dynamic Programming Valuation Model . . . . .	33
3.3.1 Formulation of the Tolling Contract Valuation . . . . .	36
3.3.2 Specification of the Underlying Commodity Price Processes . . . . .	40
3.4 Least Square Monte Carlo Method . . . . .	41
3.4.1 Commodity Price Process Simulation . . . . .	42
3.4.2 Forward Curve Modeling . . . . .	43

3.4.3	Value Function Approximation by Least-squares Regression . . . .	43
3.5	Adaptive Policy Estimation Method . . . . .	45
3.5.1	Policy Function . . . . .	45
3.5.2	The Adaptive Policy Search Algorithm . . . . .	46
3.6	Hedging Tolling Contracts . . . . .	48
3.7	Numerical Examples and Managerial Insights . . . . .	50
3.8	Conclusion and Future Work . . . . .	59
<b>IV</b>	<b>THE CONDITIONAL VALUE-AT-RISK UNDER A GARCH PRICE MODEL WITH HEAVY TAILED INNOVATIONS . . . . .</b>	<b>65</b>
4.1	Introduction . . . . .	65
4.2	Model Specification and Estimation Methodology . . . . .	67
4.2.1	Point Estimation . . . . .	67
4.2.2	Interval Estimation . . . . .	70
4.3	Simulation Study and Application . . . . .	72
4.3.1	Simulation Study . . . . .	72
4.3.2	The Bootstrap Method . . . . .	73
4.3.3	Application . . . . .	73
<b>V</b>	<b>CONCLUSION AND FUTURE RESEARCH . . . . .</b>	<b>81</b>
	<b>APPENDIX A — PROOFS IN CHAPTER 4 . . . . .</b>	<b>83</b>
	<b>REFERENCES . . . . .</b>	<b>88</b>

## LIST OF TABLES

Table 1	GMM estimations of the MRVG process . . . . .	19
Table 2	MCMC estimations of the MRVG process . . . . .	22
Table 3	The first conditional moment computed by the closed form and simulation. $\mu = 3.4657$ , $\alpha = 0.7$ , $\sigma = 0.5$ , $\nu = 0.03$ and $Y_0 = 3.4012$ are used in the simulation study . . . . .	23
Table 4	Parameters for a Hypothetical Natural Gas Fired Power Plant . . . . .	51
Table 5	Parameters for the Power Price Process . . . . .	52
Table 6	Parameters for the Gas Price Process . . . . .	53
Table 7	Estimated Values (in \$ millions) of the Tolling Agreement for Different Price Models and Heat Rates (Max_Restarts N: 3 and 6; STD: standard error). . . . .	54
Table 8	Statistics of Hedging Errors for Case 2 (Max_Restarts N: 3 and 6) . . . .	58
Table 9	Comparison of valuation by least-square Monte Carlo and adaptive policy search. (All values are in \$ millions) . . . . .	60
Table 10	Coverage probabilities by the bootstrap method for 99% one day ahead conditional VaR estimation. Nominal coverage probability level is 90%. The number of simulations is 5,000 . . . . .	73
Table 11	Coverage probabilities by the bootstrap method for 50% one day ahead conditional VaR estimation. Nominal coverage probability level is 90%. The number of simulations is 5,000 . . . . .	73

# LIST OF FIGURES

Figure 1	Real time electricity locational marginal price in the Pennsylvania - New Jersey - Maryland power market from April 1998 to September 2003. . .	6
Figure 2	One simulation path of the MRVG process with $\alpha = 0.5$ , $\mu = \log(35)$ , $\sigma = 0.5$ and $\nu = 0.001$ . . . . .	18
Figure 3	Forward curves (Contango). $\mu = 3.4657$ , $\alpha = 0.7$ , $\sigma = 0.5$ , $\nu = 0.03$ and $Y_0 = 3.4012$ . . . . .	24
Figure 4	Percentage errors of forward prices. $\mu = 3.4657$ , $\alpha = 0.7$ , $\sigma = 0.5$ , $\nu = 0.03$ and $Y_0 = 3.4012$ . . . . .	25
Figure 5	Option pricing errors (cent). The line with square markers is for call options and the line with star markers for put options. . . . .	27
Figure 6	Option pricing errors (%). The line with square markers is for call options and the line with star markers for put options. . . . .	28
Figure 7	State Transition Diagram of a Tolling Contract with the Restart Constraint	34
Figure 8	Simulated Sample Paths of Different Electricity and Gas Price Models . .	52
Figure 9	Optimal boundaries for heat rate 7.5 . . . . .	56
Figure 10	Optimal boundaries for heat rate 13.5 . . . . .	57
Figure 11	Histogram of Hedge Errors: Hr = 7.5 (left) vs. Hr = 13.5 (Right) . . . .	59
Figure 12	Estimated policies for 3 restarts remaining . . . . .	61
Figure 13	Estimated policies for 6 restarts remaining . . . . .	62
Figure 14	Policy slopes and interceptions as functions of time for 3 restarts remaining	63
Figure 15	Policy slopes and interceptions as functions of time for 6 restarts remaining	64
Figure 16	Coverage probabilities for $t(3)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction $k = 20, 22, \dots, 120$ for Student $t$ -distribution with degrees of freedom $d = 3$ . . . . .	75
Figure 17	Coverage probabilities for $t(5)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction $k = 20, 22, \dots, 120$ for Student $t$ -distribution with degrees of freedom $d = 5$ . . . . .	76
Figure 18	Coverage probabilities for $t(7)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction $k = 20, 22, \dots, 120$ for Student $t$ -distribution with degrees of freedom $d = 7$ . . . . .	77



Figure 19	Coverage probabilities for $t(9)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction $k = 20, 22, \dots, 120$ for Student $t$ -distribution with degrees of freedom $d = 9$ . . . . .	78
Figure 20	The log returns of real time electricity locational marginal price in the Pennsylvania - New Jersey - Maryland power market from April 1998 to September 2003. . . . .	79
Figure 21	Interval estimation. The estimate $\hat{x}_{\alpha,n}^t$ (broken line) and the endpoints (dotted lines) of its 90% confidence intervals are plotted against the log return $X_{t+1}$ (solid line) of PJM real time locational marginal price. I mark those points $X_{t+1}$ such that $X_{t+1} < \hat{x}_{\alpha,n}^t$ but greater than the left endpoint of the confidence interval $I_\beta^n$ with squares. . . . .	80

## SUMMARY

Restructuring the electric power industry has become a global trend. The traditional, vertically integrated electricity industry has been restructured into three separate industrial segments: generation, transmission, and distribution; and since the 1990s, competitive electricity markets have emerged. The contribution of this work is that I develop and apply financial economic methodologies to address issues in electricity spot price modeling, market-based valuation of structured electricity contracts and estimation of risk measures, and provide solutions to the problems faced by market players participating in the restructured electricity markets.

Electricity prices in competitive markets are extremely volatile with salient features such as mean-reversion and jumps and spikes. Modeling electricity spot prices is essential for asset and project valuation as well as risk management. I introduce the mean-reversion feature into a classical variance gamma model to model the electricity price dynamics as a mean-reverting variance gamma (MRVG) process. The density function and first four moments of the conditional distribution of a MRVG process are obtained. These are utilized in deriving derivative pricing formula and establishing a generalized method of moments (GMM) framework for parameter estimation.

Under a realistic electricity price model with mean-reversion and jumps, the problem of pricing and hedging electricity financial instruments is very challenging even for cases with standard electricity options, let alone for cases with the exotic options and structured transactions in which operational characteristics often need to be considered. While customized electric power contracts catering to specific business and risk management needs have gained increasing popularity among large energy firms, how to price and hedge these complex power contracts become an import and pressing issue. A tolling agreement (or tolling contract) is one such example in which a contract buyer reserves the right to take the output of an underlying electricity generation asset by paying a predetermined premium

to the asset owner. I propose a real options approach to value a tolling contract incorporating operational characteristics of the generation asset and contractual constraints. Two simulation-based methods are proposed to solve the valuation problem. The effects of different electricity price assumptions on the valuation of tolling contracts are examined. Based on the valuation model, I also propose a heuristic scheme for hedging tolling contracts and demonstrate the validity of the hedging scheme through numerical examples.

Faced with a variety of risks arising from volatile electricity prices, scarce transmission network capacity, and inelastic power demand, energy firms recognize that risk management is crucial to their business success. In particular, a firm needs to have a proper way to model and measure market price risk associated with its operations. Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized ARCH (GARCH) models are widely used to model price volatility in financial markets. Considering a GARCH model with heavy-tailed innovations for electricity price, I characterize the limiting distribution of a Value-at-Risk (VaR) estimator of the conditional electricity price distribution, which corresponds to the extremal quantile of the conditional distribution of the GARCH price process. I propose two methods, the normal approximation method and the data tilting method, for constructing confidence intervals for the conditional VaR estimator and assess their accuracies by simulation studies. I also implement the bootstrap method described in Christoffersen and Goncalves [17] for the purpose of comparison. Finally, I apply the proposed approach to electricity spot price data taken from the Pennsylvania-New Jersey-Maryland market to obtain confidence intervals of the empirically estimated Value-at-Risk of electricity prices.

# CHAPTER I

## OVERVIEW

The electricity supply industry has traditionally been a vertically integrated industry that consists of generation, transmission and distribution segments. The economic inefficiencies of this vertically integrated industry structure is well-known. Gilbert and Henly [41] show that the annual welfare losses related to the inefficiencies in the electricity industry in the United States are substantial. With the rapid technology advancement in power generation and telecommunication, reforming the traditional electricity industry to improve the economic efficiency has become viable. As a consequence, competitive electric power markets have been established in many countries, including the United Kingdom, Australia, Chile, Argentina, New Zealand, Norway and the United States, to achieve this goal.

The worldwide restructuring of the electricity industry has created three separate industrial segments: generation, transmission and distribution. Firms in different segments possess distinct risk profiles. For instance, independent power producers in the generation segment face potential risks in both revenue and production cost since they are subject to the market price risk of multiple underlying commodities (e.g. electricity and input fuel). On the other hand, utility companies, which become more focused in the business of local transmission and distribution in the restructuring process, are mostly concerned with having ample electricity supply to serve their customers at a profitable margin.

Risk hedging by a corporation should be motivated in principle by the goal of maximizing the firm's value. Hedging achieves value enhancement by reducing the likelihood of financial distress and its ensuing costs, or by reducing the variance of taxable incomes and its associated present value of future tax liabilities. Regulatory rules also play an important role in hedging practices. In California, for instance, the regulators granted the incumbent investor-owned utilities (IOUs) a fixed time frame to recover their stranded generation investments through the Competition Transition Charge. Fearing adverse market

conditions causing insufficient recovery of the stranded costs, one major utility company hired investment bankers to structure and implement an extensive hedging strategy for its stranded-cost recovery. On the other hand, the reluctance of the regulators in California to immunize the IOUs against ex-post prudence review of long-term supply contracts discouraged the adoption of such contracts, resulting in over reliance of the IOUs on the spot market for electricity procurement. This excessive exposure led to the near collapse of the California utility industry in 2001, with devastating economic losses due to prolonged outages and substantial rate increases.

As the competitive but volatile electricity markets mature, generation companies, power marketers and load serving entities (LSEs) seek certainty in their costs and revenues through hedging practices and contracting and active trading. Such activities involve quantifying, monitoring and controlling trading risks in the wholesale and retail power markets, which in turn require appropriate risk management tools and methodology.

On the supply side, managing risk associated with long-term investment in generation and transmission requires methods and tools for planning under uncertainty and for asset valuation. Many of the demands for generation asset valuation methods were spurred by the mandatory divestiture of generation assets already owned by major utility companies in various jurisdictions. For example, in California, most of the fossil-fuel plants held by the three IOUs, which account for about 60% of the total installed capacity in California by 2000, have been or will be divested to other parties. The need for asset valuation also rises from analysis of investment in new generation capacity and from efforts by regulators in the United States and abroad to develop incentives for investment in generation capacity to meet supply adequacy and system reliability objectives.

A fundamental vision underlying the worldwide movement toward a competitive electricity industry has been that most of the efficiency gains from restructuring come from long term investments in generating capacity. Under the state-ownership or required rate-of-return regulatory regime, utility companies were allowed to earn a regulated rate of return above their capital cost. Once regulators approved the construction costs of a power generating plant, the costs would be passed onto consumers through regulated electricity prices

over the life of the investment, independent of the fluctuation in market value of the investment over time due to changing energy prices, improving technology, and evolving supply and demand conditions. Most of the investment risks in generating capacity were allocated to consumers rather than producers. Firms, therefore, had few incentives to avoid excessive cost of investment and they focused on improving and maintaining quality of service rather than on developing and adopting new generation technology.

Electricity market reforms around the world have shifted much of the investment risk from consumers to producers. Under the ideal theoretical paradigm, shareholders bear all the investment risk and consumers bear the price risk, with competitive entry pushing generation capacity toward desired long-term equilibrium. In such an ideal market environment, suppliers and consumers are free to choose their desired level of risk exposure, achieved through voluntary risk management practices. Unfortunately, this idealized vision of a competitive electricity market is not working as expected, primarily due to such market imperfections as lack of demand response, abuse of locational market power, and political resistance to high prices reflecting scarcity rents and shortages.

Many political and economic issues arise as a result of the various approaches being adopted in different countries in order to restructure the electricity industry. Among these political and economic issues, I am particularly interested in electricity spot prices modeling, power contracts/assets valuation and risk management. In the remainder of this chapter, I provide a brief overview on these three aspects. In section 1.1, I outline the salient features of electricity spot prices and explain the needs and complexities in modeling electricity spot prices. The demands and importance of market-based valuation are discussed in section 1.2. In section 1.3, I introduce the application of GARCH models in practice in financial industries and the role of value-at-risk in risk management.

## ***1.1 Electricity Spot Prices Modeling***

The deregulation of the power industry has given way to a global trend toward the commoditization of electric energy. Electricity has transformed from a primarily technical business, to one in which the product is treated in much the same way as any other commodity, with

trading and risk management as key tools to run a successful business. However, we have to bear in mind that electricity is a unique commodity in that it cannot be economically stored. Moreover, the operation of an interconnected electric power network is very complicated. Therefore, most of the electricity markets are forward markets, either day-ahead or hour-ahead, complemented by real-time markets. The term “spot price” refers to not only the real-time spot prices, but also a wide range of market prices, such as the day-ahead or hour-ahead forward prices. Under the new regime of the electricity industry, the portfolios of physical assets and various supply contracts held by power marketers are exposed to market price risk. To trade electricity, perform risk management, evaluate assets and finance new investments, people have to have an in-depth understanding of electricity price behaviors. Thus, sophisticated models are needed to model electricity spot prices.

Unlike other commodities, electricity is non-storable. Furthermore, the aggregated electricity supply and demand has to be balanced continuously in order to prevent the power networks from collapsing. Since the supply and demand shocks cannot be smoothed by inventories, electricity spot prices are the most volatile among all commodity prices. Uncontrolled exposure to market price risks could lead to devastating consequences. During the summer of 1998, wholesale power prices in the Midwest of United States surged to a stunning \$7,000 per MWh from the normal price range of \$30-\$60 per MWh, causing the default of two power marketers on the East Coast. In February 2004, persistent high prices in Texas during a three-day ice storm led to the bankruptcy of a retail energy provider that was exposed to spot market prices. And of course, the California electricity crisis of 2000/2001 and its devastating economic consequences are largely attributed to the fact that the major utilities were not properly hedged through long-term supply contracts. Such expensive lessons have raised the awareness of market participants to the importance and necessity of risk management practices in a competitive electricity market. Besides the large volatility, electricity spot prices exhibit additional features.

The first such feature is seasonality. The seasonal character of electricity spot prices is a direct consequence of the fluctuations in demand. These mostly arise due to changing climate conditions, such as temperature and the number of daylight hours.

Mean-reversion, as a common feature in many traded commodities prices, also characterizes electricity spot prices. If the price of electricity is high, the supply tends to increase and thus there is a downward pressure on the price. When the power price is low, the supply of electricity tends to decrease, thus pushing the price higher.

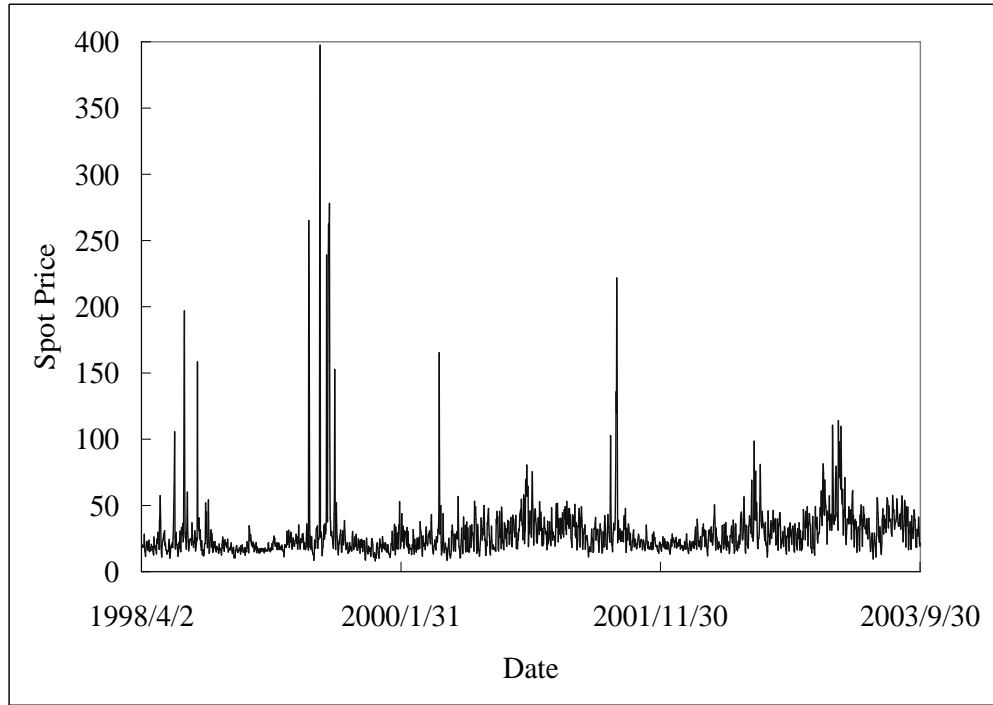
Jumps and spikes in the price process are another salient feature of electricity spot prices. Figure 1 plots the historical daily average of electricity spot prices in the western hub of the Pennsylvania - New Jersey - Maryland (PJM) power market. The jump behavior of electricity spot prices is primarily due to the fact that a typical regional aggregate supply cost curve for electricity almost always has a kink at a certain capacity level and a steep upward slope beyond that capacity level. A forced outage of a major power plant or a sudden surge in demand will either shift the supply curve left or lift up the demand curve so that the regional electricity demand curve crosses the regional supply curve at its steep-rise portion thus causing a jump in the price. When the contingency disappears in the short term, the high price will fall to its normal range, thus forming a spike.

In chapter 1, I extend the so-called Variance Gamma process to capture the above salient features of electricity spot prices. I provide the specification of the pure jump process, derive the conditional density function and moment conditions, propose two methods to estimate model parameters, and demonstrate how the prices of electricity derivatives can be obtained by transform analysis.

## ***1.2 Market-Based Valuation***

In the early 2000s, the rise and fall of the several large U.S. electric power merchants created turmoil in the power markets, and consequently caused sizable financial losses to the major financial institutions that offered loans to finance these power marketers' investment projects and business transactions. Basically, a large portion of the acquired power generation assets and the signed power purchasing contracts by the power merchants turned out to be far less profitable than what was expected due to optimistic valuations and insufficient risk management. The power marketers were unable to pay back their loans in due time and faced great financial distress. These adverse events have demonstrated the importance of





**Figure 1:** Real time electricity locational marginal price in the Pennsylvania - New Jersey - Maryland power market from April 1998 to September 2003.

an appropriate valuation and effective risk management methodology in power markets for both market participants and the financial institutions, such as banks, that have business dealings with these market participants.

Noting the extremely high price volatility, power market participants are especially wary of the price risk associated with business transactions, so they resort to customized (most likely long-term) business transactions to hedge their respective unique risk profiles, thus making the bilateral and multilateral power supply contracts ubiquitous. A market-based valuation approach is essential for pricing and risk managing these bilateral (sometimes multilateral) power transactions.

The valuation of electricity contracts differs from that of other financial contracts in that:

- a) the underlying electricity cannot be bought and sold;
- b) electricity contracts often contain

side constraints (e.g., various contract provisions) on how financial payouts are derived from the underlying electricity or a physical asset generating electricity. While a market-based valuation can be carried out by taking the price of electricity as a state variable and adopting a proper discounting factor, these side-constraints significantly increase the complexity of pricing electricity contracts. I propose a market-based approach for pricing and hedging electricity contracts with a complex contractual structure. I outline typical operational and contractual provisions in a structured electricity supply contract and incorporate them into a real options valuation framework. This approach is a valuable tool for both power market participants and those financial institutions that are interested in exploring business opportunities in power markets.

The discounted cash flow method (DCF) was the norm for valuing power supply contracts and evaluating generation/transmission asset investments in the traditionally regulated electricity industry, as power price was set by regulators based on cost of service. The basis of DCF valuation is a set of static (or estimated) future cash flow. However, the electricity prices are no longer preset in the newly restructured power industry. They are driven by the ever-changing fundamental market supply and demand conditions. Under the new regime, a DCF valuation approach, which is based on static cash flow estimates rather than a dynamically evolving cash flow, undervalues power contracts and assets because it fails to capture the value associated with the inherent optionality for dynamically maximizing the cash flow of an underlying asset and takes little account of the extraordinary electricity price volatility into the valuation. Deng et al. [25] propose a *real options* approach based on an analogy between the payoff of certain financial options and that of a physical asset for power asset valuation. They demonstrate that the option-pricing approach is the better alternative to the DCF method based on market information. Deng and Oren [28] and Tseng and Barz [90] advance the real options valuation of power plants further by incorporating operational constraints into the valuation framework.

I will present two simulation-based methods to value a tolling agreement with some physical and contractual constraints in the real option valuation framework. One method is to solve a dynamic programming problem with least square Monte Carlo technique and the

other is subject to the estimation of a policy function which specifies the contract execution policy. While simulating price data, I make use of market forward information such as a forward price curve and implied volatilities. I also propose a heuristic delta hedging strategy in chapter 3.

### ***1.3 GARCH Models and Value at Risk***

Two important empirical features about financial return series have drawn considerable attention in the field of financial econometrics, namely, heteroscedasticity and heavy-tailed phenomenon. For example, the recent Séminaire Européen de Statistique reported in Finkenzstädt and Rootzén (2004) consists of excellent reviewed articles on a variety of research topics related to these two features. As an attempt for capturing these stylized empirical findings in financial data, ARCH and generalized ARCH (GARCH) models were proposed to explicitly model the conditional second moments and their long-range dependence structure. The classical ARCH/GARCH models are based on conditional Gaussian innovations (see Engle [36] and Bollerslev [7]). They can be used to model risk attributes such as volatility clustering and the long-range dependence structure that exists in equity prices, financial indices, and foreign exchange rates (see Bollerslev et al. [8] and Taylor [87]).

There is growing literature on applications of ARCH/GARCH models in asset pricing and risk management. With ubiquitous risks in financial markets, one of the most important tasks of financial institutions is to evaluate the exposure to market risks. This is commonly done by estimating the so-called Value-at-Risk (VaR). Market risks experienced during extreme market movements can cause dramatic changes in portfolio values. This can create huge profits or losses for financial institutions and may lead to financial pitfalls as demonstrated in the Long Term Capital Management case. VaR measures market risks by providing a single estimate of the worst possible financial loss to a portfolio over a fixed time horizon for a given confidence (or probability) level (see Jorion [57], Rachev and Mitnik [71] and Duffie and Pan [34] for a general introduction and exposition of VaR). Mathematically, VaR is defined as a quantile of a probability distribution, which is used to model an underlying portfolio value or its return. Financial institutions and regulators

use VaR to quantify market risks and set capital reserves for market risks. For instance, traders at financial institutions often have their trading limits specified in terms of daily VaR of their trading books. Another appealing implication of VaR is that it can be utilized as a vehicle for corporate self-insurance since VaR can be interpreted as the amount of uninsured loss acceptable to a corporation (see Shimko [86]). A corporation should buy external insurance when the self-insurance losses, as reflected by VaR measures, are greater than the cost of insurance by hedging.

In practice, a key risk measure for financial institutions based on the VaR concept is the conditional VaR, which is the worst possible loss due to adverse market movements over the next reporting period (e.g., a day or a week) *conditional* on current portfolio volatility and market information. This quantity corresponds to the tails of the conditional profit-and-loss (P&L) distribution of a portfolio. It is essentially the basis for setting portions of the day-to-day operating capital reserves for many financial institutions. As the GARCH models have been successfully applied in modeling the P&L distribution and the volatility structure of a portfolio of securities and other financial assets, the conditional VaR of a GARCH model becomes an important quantity to study. An additional important trait of the conditional VaR is the robustness property of the conditional VaR estimator. When financial institutions utilize conditional VaR for setting capital reserves, they first need to estimate it based on some statistical models, either parametrically or non-parametrically.

Empirical evidence has demonstrated that the conditional normal time series models (e.g., the classical GARCH models) are inadequate in estimating the tail quantiles of conditional return distributions (see Danielsson and de Vries [21] for instance). This prompts the gradual adoption of models with heavy-tailed innovations in risk modeling practice. Many extensions of the classical GARCH models with heavy-tailed innovations have been proposed. McNeil and Frey [68] consider a GARCH model with generalized Pareto distributed innovations and propose a two-step approach to estimate the conditional VaR. While their idea seems intuitive, important statistical properties such as confidence interval estimation and asymptotic properties remain largely unexplored.

In chapter 4, I first derive the limiting distribution of the extreme conditional VaR

estimator in McNeil and Frey [68]. Instead of working within the framework of generalized Pareto distribution as in McNeil and Frey, I deal with the heavy-tailed innovations. I construct confidence intervals of a conditional VaR estimator developed from the extreme value theory by the traditional normal approximation method based on the asymptotic normality of the VaR estimator and the recent data tilting method studied in Hall and Yao [46] and Peng and Qi [79].

## ***1.4 Contributions of this Work***

In this thesis, the goal is to develop and apply financial economic methodologies to address issues in electricity spot price modeling, market-based valuation and risk management. I attempt to provide solutions to the problems faced by market players participating in the restructured electricity markets.

In chapter 2, a new stochastic process is constructed to model electricity spot prices, which can capture the realistic aspects of electricity prices. For the proposed electricity price model, I derive the conditional density function and moment conditions, propose two methods to estimate model parameters and provide a framework for pricing electricity derivatives.

A market-based valuation framework for customized electricity supply contracts with a complex structure is proposed in chapter 3. I propose two simulation-based real options methods to value a tolling agreement with multiple exercising decisions. Moreover, recognizing the fact that there are often transaction costs incurred when exercising the embedded options of a tolling agreement due to the operational and contractual constraints, I further them into the valuation problem. I first discretize the stochastic processes for the underlying commodity prices and then formulate the valuation as a stochastic dynamic programming (SDP) problem. The least square Monte Carlo method is employed to solve the SDP. The second method for the tolling agreement valuation problem is subject to estimating an exercise policy function whose value is the basis for making an operational decision. According to the empirical evidence, I propose a piecewise linear policy function that is a function of time and state variables. Parameters in the policy function are estimated in a subspace of

the parameter space. While the fitted policy function is sub-optimal, numerical experiments demonstrate that the resulting contract values match closely with those obtained by the first method. I further propose a heuristic delta hedging strategy for buyers of a tolling agreement to hedge their price risks. The heuristic hedging strategy works well in general, particularly better for tolling contracts associated with efficient power plants.

In chapter 4, I estimate the conditional Value-at-Risk (VaR) based on GARCH(1,1) model with heavy tailed innovations and further derive the limiting distribution of the extreme conditional VaR estimator. Moreover, I construct the confidence interval of the conditional Value-at-Risk estimator by a normal approximation method and a data tilting method. The knowledge of the confidence interval of the conditional VaR can be highly valuable in applications such as setting prudent capital reserve requirements for banks and conservative trading limits for traders as well as evaluating corporate self-insurance exposures. It provides upper and lower bounds of the VaR estimator at a certain confidence level rather than a single point estimate.

Finally in chapter 5, I summarize my thesis and point out several directions which deserve further investigation for future research.

## CHAPTER II

### ELECTRICITY SPOT PRICES MODELLING

#### *2.1 Introduction*

As noted in section 1.1, modeling electricity spot prices is a very challenging task for researchers because of the distinguishing characteristics of electricity. First of all, electricity is almost non-storable. Moreover, the aggregated electricity supply and demand has to be balanced continuously so as to prevent the electric power networks from collapsing. Therefore, electricity spot prices are extremely volatile with the salient features, such as mean-reversion and the presence of jumps and spikes.

There have been some studies on modeling electricity spot prices. Generally, electricity price models are classified in four groups: production costs models, time series models, continuous time stochastic models and game theory models.

Fundamental models based on production cost are developed for centralized systems. These models simulate the operation of power systems and consider not only production cost but also the agents' strategic behavior impact on market price. This group includes Batlle [2] and Schweppe et al. [85].

Time series models study price evolution from a statistical point of view. The well-known generalized autoregressive conditional heteroscedastic (GARCH) models suggested by Bollerslev in [7] are representative in this group. A multivariate version of the GARCH model has been applied to generate fuel price scenarios for risk analysis in a wholesale electricity market in Batlle and Barquín [3]. Worthington et al. [95] examine the transmission of spot electricity prices and price volatility among the five regional electricity markets in the Australian National Electricity Market and use a multivariate GARCH model to identify the source and magnitude of price and price volatility spillovers. Longstaff and Wang [64] employ a GARCH(1,1) model to estimate the conditional variance of unexpected electricity price changes and use the GARCH estimate as the ex ante price risk measure to estimate

the forward premium. Other studies in this group include Goto and Karolyi [43], Duffie and Gray [33], Bunn [11] and Mount and Ethier [72].

Continuous time stochastic models of commodity prices first come from economic and financial world due to the success of geometric Brownian motion in modeling stock prices (see Black and Scholes [6]). Early studies in this area typically assume that commodity prices are governed by geometric Brownian motion. More recently, a number of researchers have considered the use of mean-reverting price models such as Knittel and Roberts [61], Deng et al. [25], Lucia and Schwartz [65] and Tseng and Barz [90]. The basic idea is that the deviations of the price from its equilibrium level are corrected and subjected to random perturbations. Schwartz and Smith [82] propose a two-factor mean-reverting model. The underlying idea is that the short-term deviations correspond to temporary changes in prices that are expected to revert toward zero, and changes in the equilibrium price level reflect fundamental longer term changes that are expected to persist, such as expectations of the exhaustion of existing supply, improving technology for the production and discovery of the commodity, and inflation, as well as political and regulatory effects. Kaminski [58] has pointed out the necessity of introducing jumps and stochastic volatility in modeling electricity prices. Barz and Johnson [1] pointed out the inadequacy of the Geometric Brownian motion and mean-reverting process in modeling electricity prices. Deng [24] examined three types of mean-reverting jump-diffusion electricity price models: mean-reverting jump-diffusion process with deterministic volatility, mean-reverting jump-diffusion process with regime-switching, and mean-reverting jump-diffusion process with stochastic volatility. Recently, González et al. [42] propose an input-output hidden Markov model to analyze and forecast electricity prices. This model is based on artificial neural networks (ANNs) (see Hornik et al. [55]). In their model, they consider different market states. Each market state is characterized by a particular density function, which represents the relationship between explanatory variables such as load, hydro, thermal and nuclear resources and the electricity spot price through a dynamic regression model. Hence, their model can be seen as a switching model in which the system evolves through different states, where a particular dynamic regression model is adjusted in each one.



Another group of models is to obtain reasonable medium-term price estimations and analysis market power based on game theory. Equilibrium models described in Ventosa et al. [92] take the analysis of strategic market equilibrium (the market reaches equilibrium when each firm's strategy is the best response to the strategies actually employed by its opponents). The concept of Nash equilibrium is the foundation of such models.

In this chapter, a new continuous time stochastic process is proposed to model electricity spot price. The idea here rises from the so-called variance gamma (VG) process. The new stochastic process is constructed in a similar way as the variance gamma process, but introduces the mean-reverting component. I expect that the new process, which we call the mean-reverting variance gamma (MRVG) process, can capture the noticeable features such as mean-reversion and jumps and spikes. The conditional density function at time  $t$  can be given in an integral form, and the first four conditional moments can be explicitly expressed in terms of model parameters. Parameter estimation is performed by two classical methods: the generalized method of moments and the Markov chain Monte Carlo method. Derivatives pricing formulae are obtained through transform analysis.

## 2.2 *Variance Gamma Process Review*

The variance gamma process, proposed by Madan et al. [66], is a three parameter generalization of Brownian motion. The VG process can be obtained by evaluating Brownian motion with constant drift and volatility at a random time change given by a gamma process. Each unit of calendar time may be viewed as having an economically relevant time length given by a gamma random variable with unit mean and positive variance. Unlike most existing price models, the VG process has no continuous martingale component. Instead, it is a pure jump process with an infinite arrival rate. High activity (this may loosely be measured by the volume or number of transactions) is accounted for by the infinite number of jumps in any interval of time. Hence, there is no need to introduce an additional diffusion component.

Let

$$b(t; \theta, \sigma) = \theta t + \sigma W(t) \tag{1}$$

where  $W(t)$  is a standard Brownian motion. The process  $b(t; \theta, \sigma)$  is a Brownian motion

with drift  $\theta$  and volatility  $\sigma$ .

The gamma process  $\gamma(t; \mu, \nu)$  with mean rate  $\mu$  and variance rate  $\nu$  is a stochastic process with independent and stationary increments having gamma distributions. The density of the increment  $g = \gamma(t+h; \mu, \nu) - \gamma(t; \mu, \nu)$  over the time interval  $(t, t+h)$  is given by the gamma density function with mean  $\mu h$  and variance  $\nu h$  as follows:

$$f_h(g) = \left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 h}{\nu}} \frac{g^{\frac{\mu^2 h}{\nu}-1} e^{-\frac{\mu}{\nu} g}}{\Gamma(\frac{\mu^2 h}{\nu})}, g > 0 \quad (2)$$

where  $\Gamma(x)$  is the gamma function. The characteristic function of the gamma density  $\Phi_{\gamma(t)}(u) = E[\exp(iu\gamma(t; \mu, \nu))]$  is given by:

$$\Phi_{\gamma(t)}(u) = \left(\frac{1}{1 - iu\frac{\nu}{\mu}}\right)^{\frac{\mu^2 t}{\nu}} \quad (3)$$

The VG process  $X(t; \sigma, \nu, \theta)$  is defined in terms of a Brownian motion with drift  $b(t; \theta, \sigma)$  and a gamma process with unit mean rate  $\gamma(t; 1, \nu)$  by evaluating the Brownian motion at a gamma time change.

$$X(t; \sigma, \nu, \theta) = b(\gamma(t; 1, \nu); \theta, \sigma) \quad (4)$$

Hence, the density function for the VG process at time  $t$  conditional on the realization of the gamma time change  $g$  can be expressed as a normal density. The unconditional density function then can be obtained by integrating out  $g$  with respect to the gamma density (2). This gives us  $f_{X(t)}(X)$  as

$$f_{X(t)}(X) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi g}} \exp\left(-\frac{(X - \theta g)^2}{2\sigma^2 g}\right) \frac{g^{\frac{t}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg \quad (5)$$

The characteristic function for the VG process  $\Phi_{X(t)}(u) = E[\exp(iuX(t; \sigma, \nu, \theta))]$  is thus given by

$$\Phi_{X(t)}(u) = \left(\frac{1}{1 - iu\theta\nu + (\sigma^2\nu/2)u^2}\right)^{t/\nu} \quad (6)$$

Unlike Brownian motion, the VG process is a process of finite variation and thus can be written as the difference of two increasing processes. In the case of the VG process, it can be expressed as the difference of two independent increasing gamma processes as follows:

$$X(t; \sigma, \nu, \theta) = \gamma_p(t; \mu_p, \nu_p) - \gamma_n(t; \mu_n, \nu_n) \quad (7)$$

where  $\mu_p, \nu_p, \mu_n, \nu_n$  have explicit relations with the original parameters of the VG process in (4) (see Madan et al. [66]).  $\gamma_p(t)$  and  $\gamma_n(t)$  account for the market up and down moves respectively. Explicit expressions for the first four central moments over an interval of length  $t$  are also derived in Madan et al. [66] and given as follows:

$$\begin{aligned}
E[X(t)] &= \theta t \\
E[(X(t) - E[X(t)])^2] &= (\theta^2 \nu + \sigma^2) t \\
E[(X(t) - E[X(t)])^3] &= (2\theta^3 \nu^2 + 3\sigma^2 \theta \nu) t \\
E[(X(t) - E[X(t)])^4] &= (3\sigma^4 \nu + 12\sigma^2 \theta^2 \nu^2 + 6\theta^4 \nu^3) t \\
&\quad + (3\sigma^4 + 6\sigma^2 \theta^2 \nu + 3\theta^4 \nu^2) t^2
\end{aligned} \tag{8}$$

### 2.3 Mean-Reverting Variance Gamma Process

As a model nesting the Black Scholes model as a parametric special case, the VG process has successfully modeled the dynamics of log stock prices/indices (See Madan et al. [66]). Unfortunately, due to the absence of a mean-reversion component, the VG process is not an appropriate candidate to model the dynamics of electricity spot prices. To capture the mean-reversion feature of electricity prices, I extend the VG process by evaluating a simple mean-reverting process at a gamma time.

Let  $\{S(t) : t \geq 0\}$  denote a mean-reverting process governed by the following stochastic differential equation:

$$dS = \alpha(\mu - S)dt + \sigma dW(t) \tag{9}$$

where  $W(t)$  is a standard Brownian motion,  $\alpha$  is the mean-reverting speed and  $\mu$  is the long-term mean of  $S(t)$ . Suppose  $t \geq s \geq 0$  and the information set at time  $s$  is denoted by  $\mathcal{F}_s$ , then  $S(t|\mathcal{F}_s)$  is normally distributed with mean and variance as follows:

$$S(t)|\mathcal{F}_s \sim N(\mu + (S(s) - \mu)e^{-\alpha(t-s)}, \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(t-s)})) \tag{10}$$

The mean-reverting variance gamma process  $Y(t) : t \geq 0$  then is obtained by replacing the time  $t$  in (9) by the gamma process with unit mean  $\gamma(t; 1, \nu)$  as follows:

$$dY = \alpha(\mu - Y)d\gamma(t; 1, \nu) + \sigma dW(\gamma(t; 1, \nu)) \tag{11}$$

Conditional on the realization of the gamma time change  $g = \gamma(t; 1, \nu) - \gamma(s; 1, \nu)$ ,  $0 \leq s \leq t$  and the information set  $\mathcal{F}_s$ ,  $Y(t)$  has a normal distribution with mean and variance as in (10) for the time change  $g$ . Integrating the normal density with respect to the gamma density, we can obtain the conditional density function  $f(Y(t)|\mathcal{F}_s)$  as

$$f(Y(t)|\mathcal{F}_s) = \int_0^\infty \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha g})}} e^{-\frac{(Y(t)-\mu-(Y(s)-\mu)e^{-\alpha g})^2}{\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha g})}} \frac{g^{\frac{t-s}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t-s}{\nu}} \Gamma(\frac{t-s}{\nu})} dg \quad (12)$$

The conditional  $p^{th}$  raw moment can be obtained by directly integrating out  $Y(t)^p$  with the condition density function. We derive the first four conditional moments as follows:

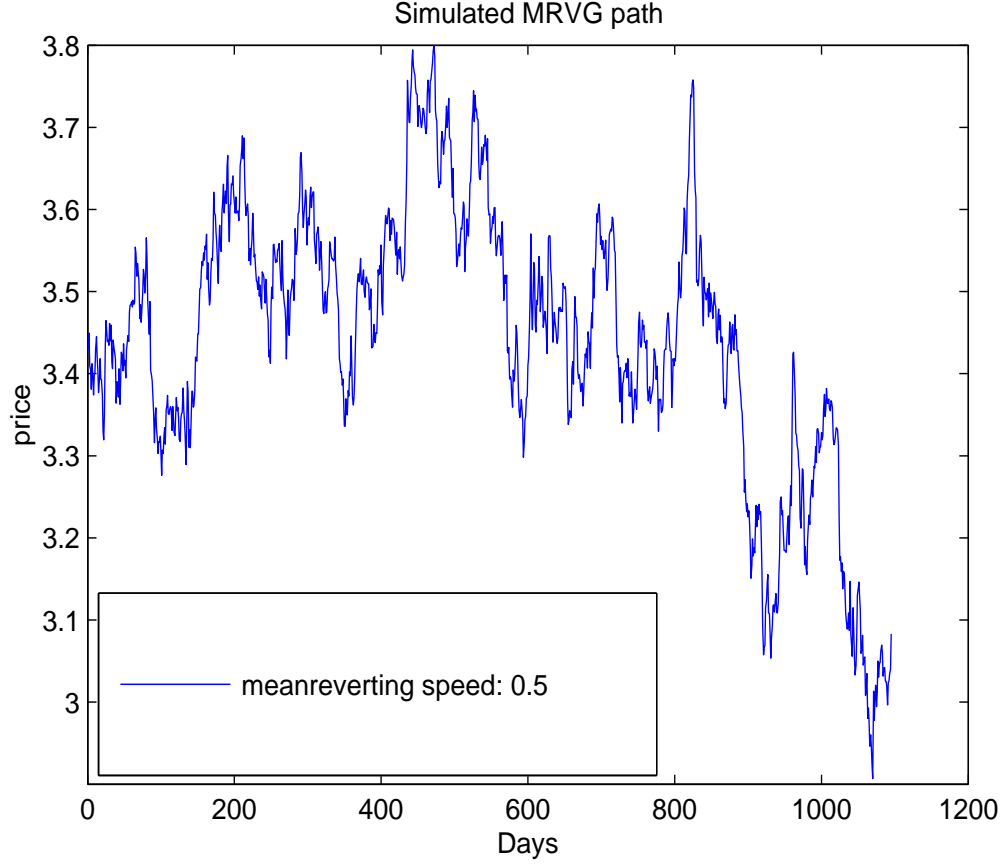
$$\begin{aligned} E[Y(t)|\mathcal{F}_s] &= \mu + \frac{Y(s)-\mu}{(1+\alpha\nu)^{\frac{t-s}{\nu}}} \\ E[Y(t)^2|\mathcal{F}_s] &= \mu^2 + \frac{\sigma^2}{2\alpha} + \frac{2\mu(Y(s)-\mu)}{(1+\alpha\nu)^{\frac{t-s}{\nu}}} + \frac{(Y(s)-\mu)^2 - \frac{\sigma^2}{2\alpha}}{(1+2\alpha\nu)^{\frac{t-s}{\nu}}} \\ E[Y(t)^3|\mathcal{F}_s] &= \mu^3 + \frac{3\mu\sigma^2}{2\alpha} + \frac{3(\mu^2 + \frac{\sigma^2}{2\alpha})(Y(s)-\mu)}{(1+\alpha\nu)^{\frac{t-s}{\nu}}} + \frac{3\mu((Y(s)-\mu)^2 - \frac{\sigma^2}{2\alpha})}{(1+2\alpha\nu)^{\frac{t-s}{\nu}}} \\ &\quad + \frac{(Y(s)-\mu)^3 - (Y(s)-\mu)\frac{3\sigma^2}{2\alpha}}{(1+3\alpha\nu)^{\frac{t-s}{\nu}}} \\ E[Y(t)^4|\mathcal{F}_s] &= \mu^4 + \frac{3\mu\sigma^2}{\alpha} + \frac{3\sigma^4}{4\alpha^2} + \frac{2\mu(2\mu^2 + \frac{3\sigma^2}{\alpha})(Y(s)-\mu)}{(1+\alpha\nu)^{\frac{t-s}{\nu}}} + \frac{3(2\mu^2 + \frac{\sigma^2}{\alpha})((Y(s)-\mu)^2 - \frac{\sigma^2}{2\alpha})}{(1+2\alpha\nu)^{\frac{t-s}{\nu}}} \\ &\quad + \frac{4\mu(Y(s)-\mu)^3 - (Y(s)-\mu)\frac{6\mu\sigma^2}{\alpha}}{(1+3\alpha\nu)^{\frac{t-s}{\nu}}} + \frac{(Y(s)-\mu)^4 - \frac{3\sigma^2}{\alpha}(Y(s)-\mu)^2 + \frac{3\sigma^4}{4\alpha^2}}{(1+4\alpha\nu)^{\frac{t-s}{\nu}}} \end{aligned} \quad (13)$$

## 2.4 Simulation Study

I perform some simulation studies on the MRVG process. Figure 2 plots one simulation path of the MRVG process. To estimate the parameters, I propose a Markov Chain Monte Carlo (MCMC) statistical estimation method. Derivatives pricing formulae are derived by the transform analysis.

### 2.4.1 Model Parameters Estimation

In this section, I implement the generalized method of moments (GMM) (see Hansen [49] and Hamilton [48]) and the hybrid Metropolis within Gibbs sampling that is in the class of Markov chain Monte Carlo (MCMC) algorithm (See Neal [74], Cheng [16] and Eraker [37]) to estimate parameters. The generalized method of moments utilizes the conditional moments condition given in (13). The parameters can be estimated by matching theoretical moments to empirical moments. Thus, an optimization procedure is needed and plays an important



**Figure 2:** One simulation path of the MRVG process with  $\alpha = 0.5$ ,  $\mu = \log(35)$ ,  $\sigma = 0.5$  and  $\nu = 0.001$ .

role in the implementation. Due to the complexity of moments conditions, the objective function is very complicated and therefore the optimization problem is hard to solve in practice. The Markov chain Monte Carlo method is an approach in the Bayesian framework. A time series whose stationary distribution is the joint posterior distribution of model parameters is generated. However, due to the non-tractable conditional density function, numerical integral technique has to be employed to calculate the density. Hence in practice, it will take a very long time to get the stationarity.

#### *2.4.1.1 The Generalized Method of Moments*

The generalized method of moments has been used for a long time to estimate parameters. The general statement of GMM was developed by Hansen(1982). The GMM estimator is

derived by minimizing a criterion function based on theoretical and sample moments as:

$$Q(\theta; Y_0, \dots, Y_T) \equiv \mathbf{g}'\mathbf{W}\mathbf{g} \quad (14)$$

where  $\mathbf{g}$  is a vector of moment conditions.  $\mathbf{W}$  is a positive definite symmetric weighting matrix reflecting the importance given to matching each of the moments.

To form a GMM estimator, I choose the following moment condition vector (see Zhou [96]):

$$h_t(\theta) = \begin{bmatrix} E(Y_{t+1}|Y_t) - Y_{t+1} \\ (E(Y_{t+1}|Y_t) - Y_{t+1})Y_t \\ E(Y_{t+1}^2|Y_t) - Y_{t+1}^2 \\ (E(Y_{t+1}^2|Y_t) - Y_{t+1}^2)Y_t \\ (E(Y_{t+1}^2|Y_t) - Y_{t+1}^2)Y_t^2 \end{bmatrix} \quad (15)$$

which satisfies the orthogonality conditions. Let  $g_T(\theta) \equiv (1/T) \sum_{t=1}^T h_t(\theta)$  the sample average of  $h_t(\theta)$ . Then the GMM estimator  $\theta$  is the value of  $\theta$  that minimizes the criterion function given in (14). During the implementation, I employ the Newey-West estimator (see Newey and West [75]) to estimate the asymptotic variance matrix  $S$  of  $g_T(\theta)$  and then the weighting matrix in (14) is the inverse of  $S$ . All estimates are reported in table 1. The true values in the table are the values I use to simulate data.

**Table 1:** GMM estimations of the MRVG process

<i>Parameters</i>	$\alpha$	$\sigma$	$\nu$	$\mu$
<i>TrueValues</i>	0.5	0.3	0.05	3.47
<i>MeanofEst.</i>	0.96	0.22	0.25	3.59
<i>STD</i>	0.75	0.10	0.34	1.39

#### 2.4.1.2 The Markov Chain Monte Carlo Method

Consider a time series of  $\{Y_t, t = 0, 1, 2, \dots, T\}$  generated according to a MRVG process. Let  $\theta = (\mu, \alpha, \sigma, \nu)$  denote the parameter vector. Under the Bayesian framework, the posterior distribution of parameters is the conditional density  $P(\theta|Y_0, Y_1, \dots, Y_T)$  which

can be expressed as a proportionality in terms of the likelihood and the prior as follows:

$$P(\theta|Y_0, Y_1, \dots, Y_T) \propto P(Y_0, Y_1, \dots, Y_T|\theta)P(\theta) \quad (16)$$

where  $P(\theta)$  is the prior and the likelihood  $P(Y_0, Y_1, \dots, Y_T|\theta)$  could be obtained as

$$\prod_{i=0}^{n-1} f(Y_{t+i\Delta}|Y_{t+(i-1)\Delta}, \theta) \quad (17)$$

where  $f(Y_{t+i\Delta}|Y_{t+(i-1)\Delta}, \theta)$  is given by (12)

The Metropolis Hasting within Gibbs sampling is implemented according to the following simulation scheme:

First, I arbitrarily choose some points  $(\mu^0, \alpha^0, \sigma^0, \nu^0)$  as starting points. For  $g = 1, \dots, G$ ,  $(\mu^g, \alpha^g, \sigma^g, \nu^g)$  are generated step by step as follows:

- $\mu^g \sim P(\mu|Y_0, Y_1, \dots, Y_T, \alpha^{g-1}, \sigma^{g-1}, \nu^{g-1})$

In this step, I generate a new point for  $\mu$  as follows:

$$\mu^* = \mu^{g-1} + s_1 * \epsilon_1 \quad (18)$$

where  $\epsilon_1 \sim N(0, 1)$ . Based on a uniform prior  $\pi(\mu) \sim U[b_1, b_2]$ , I accept the  $\mu^*$  with probability  $\omega$ :

$$\omega = \min\left[\frac{p(Y_0, Y_1, \dots, Y_T|\mu^*, \alpha^{g-1}, \sigma^{g-1}, \nu^{g-1})}{p(Y_0, Y_1, \dots, Y_T|\mu^{g-1}, \alpha^{g-1}, \sigma^{g-1}, \nu^{g-1})}, 1\right] \quad (19)$$

Then,

$$\mu^g = \begin{cases} \mu^* & \text{if } \mu^* \text{ is accepted.} \\ \mu^{g-1} & \text{if } \mu^* \text{ is rejected.} \end{cases} \quad (20)$$

- $\alpha^g \sim P(\alpha|Y_0, Y_1, \dots, Y_T, \mu^g, \sigma^{g-1}, \nu^{g-1})$

Similar to the previous step, I propose a new point for  $\alpha$  as follows:

$$\alpha^* = \alpha^{g-1} + s_2 * \epsilon_2 \quad (21)$$

where  $\epsilon_2 \sim N(0, 1)$ . With a uniform prior  $\pi(\alpha) \sim U[b_3, b_4]$ , I then accept the  $\alpha^*$  with probability  $\omega$ :

$$\omega = \min\left[\frac{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^*, \sigma^{g-1}, \nu^{g-1})}{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^{g-1}, \sigma^{g-1}, \nu^{g-1})}, 1\right] \quad (22)$$

Then,

$$\alpha^g = \begin{cases} \alpha^* & \text{if } \alpha^* \text{ is accepted.} \\ \alpha^{g-1} & \text{if } \alpha^* \text{ is rejected.} \end{cases} \quad (23)$$

- $\sigma^g \sim P(\sigma|Y_0, Y_1, \dots, Y_T, \mu^g, \alpha^g, \nu^{g-1})$

To update  $\sigma$ ,  $\sigma^*$  is proposed as follows:

$$\sigma^* = \sigma^{g-1} + s_3 * \epsilon_3 \quad (24)$$

where  $\epsilon_3 \sim N(0, 1)$ . Given a uniform prior  $\pi(\sigma) \sim U[b_5, b_6]$ , the  $\sigma^*$  is accepted with probability  $\omega$ :

$$\omega = \min\left[\frac{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^g, \sigma^*, \nu^{g-1})}{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^g, \sigma^{g-1}, \nu^{g-1})}, 1\right] \quad (25)$$

And then  $\sigma^g$  is given by

$$\sigma^g = \begin{cases} \sigma^* & \text{if } \sigma^* \text{ is accepted.} \\ \sigma^{g-1} & \text{if } \sigma^* \text{ is rejected.} \end{cases} \quad (26)$$

- $\nu^g \sim P(\nu|Y_0, Y_1, \dots, Y_T, \mu^g, \alpha^g, \sigma^g)$

To generate  $\nu^g$ , I implement the Metropolis Hasting scheme by proposing

$$\nu^* = \nu^{g-1} + s_4 * \epsilon_4 \quad (27)$$

where  $\epsilon_4 \sim N(0, 1)$ . The acceptance probability  $\omega$  for the  $\nu^*$  is:

$$\omega = \min\left[\frac{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^g, \sigma^g, \nu^*)}{p(Y_0, Y_1, \dots, Y_T|\mu^g, \alpha^g, \sigma^g, \nu^{g-1})}, 1\right] \quad (28)$$

Thus,  $\nu^g$  is:

$$\nu^g = \begin{cases} \nu^* & \text{if } \nu^* \text{ is accepted.} \\ \nu^{g-1} & \text{if } \nu^* \text{ is rejected.} \end{cases} \quad (29)$$

Table 2 reports the estimation results by the hybrid MCMC method. Due to the limitation of computational facilities, I do not generate the time series long enough to achieve the stationarity. Therefore the estimation is not so satisfactory and depends on the initial values I use.



**Table 2:** MCMC estimations of the MRVG process

<i>Parameters</i>	$\alpha$	$\sigma$	$\nu$	$\mu$
<i>TrueValues</i>	0.5	0.3	0.05	3.47
<i>MeanofEst.</i>	0.49	0.50	0.05	3.58
<i>STD</i>	0.26	0.26	0.03	1.79

### 2.4.2 Derivatives Pricing

Suppose the electricity spot price is governed by a mean-reverting variance gamma process (11) in a risk neutral world. Then the prices of European type contingent claims on the underlying electricity can be obtained by the transform analysis.

Suppose  $Y_t$  is the state variable in  $\mathbb{R}$  and  $u \in C^n$ , the generalized transform function under the risk neutral measure  $Q$  of  $Y_T$  conditional on  $\mathcal{F}_t$  when  $t < T$  is given by

$$\Psi(u, Y_t, t, T) = E^Q[e^{-r(T-t)}e^{uY_T}|\mathcal{F}_t] \quad (30)$$

let  $\Delta = T - t$ , we have

$$\begin{aligned} \Psi(u, Y, t, T) &= E^Q[e^{-r\Delta}e^{uY_T}|\mathcal{F}_t] \\ &= e^{-r\Delta} \int e^{uy_T} \left( \int_0^\infty \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha g})}} e^{-\frac{(y_T-\mu-(y_t-\mu)e^{-\alpha g})^2}{\frac{\sigma^2}{\alpha}(1-e^{-2\alpha g})}} \frac{g^{\frac{\Delta}{\nu}-1}e^{-\frac{g}{\nu}}}{\nu^{\frac{\Delta}{\nu}}\Gamma(\frac{\Delta}{\nu})} dg \right) dy_T \\ &= e^{-r\Delta} \int_0^\infty \left( \int e^{uy_T} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha g})}} e^{-\frac{(y_T-\mu-(y_t-\mu)e^{-\alpha g})^2}{\frac{\sigma^2}{\alpha}(1-e^{-2\alpha g})}} dy_T \right) \frac{g^{\frac{\Delta}{\nu}-1}e^{-\frac{g}{\nu}}}{\nu^{\frac{\Delta}{\nu}}\Gamma(\frac{\Delta}{\nu})} dg \\ &= e^{-r\Delta} \int_0^\infty e^{(\mu+(y_t-\mu)e^{-\alpha g})u + \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha g})u^2/2} * \frac{g^{\frac{\Delta}{\nu}-1}e^{-\frac{g}{\nu}}}{\nu^{\frac{\Delta}{\nu}}\Gamma(\frac{\Delta}{\nu})} dg \end{aligned} \quad (31)$$

Let  $G_{a,b}(x; Y_0, T, \theta)$  denote the price of a security that pays  $e^{a*Y_T}$  at time T in the event that  $b * Y_T \leq x$ . Then we have

$$\begin{aligned} G_{a,b}(x; Y_0, T, \theta) &= E^Q[e^{-r(T-t)}e^{a*Y_T}\mathbf{1}_{b*Y_T \leq x}|\mathcal{F}_t] \\ &= \frac{\Psi(a, Y_0, 0, T)}{2} - \frac{1}{\pi} \int_0^\infty \frac{Im[\Psi(a+ivb, Y_0, 0, T)e^{-ivx}]}{v} dv \end{aligned} \quad (32)$$

For properly chosen constants  $x, a$  and  $b$ ,  $G_{a,b}(x; Y_0, T, \theta)$  can serve as a building block in pricing contingent claims such as forward/futures and call/put options.

Table 3 reports the first conditional moment of  $Y_t$  given by the first equation in (13) and sample first moment of simulated data for 12 months. One can see that they are almost the same. Therefore, I use the derivative prices calculated based on the simulated data set as a benchmark later.

**Table 3:** The first conditional moment computed by the closed form and simulation.  $\mu = 3.4657$ ,  $\alpha = 0.7$ ,  $\sigma = 0.5$ ,  $\nu = 0.03$  and  $Y_0 = 3.4012$  are used in the simulation study

<i>Month</i>	<i>ClosedFormValue</i>	<i>SimulationValue</i>
1	3.4049	3.4049
2	3.4080	3.4081
3	3.4113	3.4114
4	3.4143	3.4145
5	3.4173	3.4174
6	3.4200	3.4201
7	3.4226	3.4228
8	3.4250	3.4252
9	3.4273	3.4275
10	3.4295	3.4297
11	3.4315	3.4317
12	3.4335	3.4337

#### 2.4.2.1 Futures/Forward Price

Electricity forward contracts represent the obligation to buy or sell a fixed amount of electricity at a pre-specified contract price, known as the forward price, at certain time in the future (called maturity or expiration time). The payoff of a forward contract promising to deliver one unit of electricity at price  $F$  at a future time  $T$  is:

$$\text{Payoff of a Forward Contract} = S_T - F \quad (33)$$

where  $S_T$  is the electricity spot price at time  $T$ . Although the payoff function (33) appears to be the same as for any financial forwards, electricity forwards differ from other financial and commodity forward contracts in that the underlying electricity is a different commodity at different times. Since entering into a forward contract is free, under a risk neutral measure,

the forward price  $F$  at time  $t$  is given by:

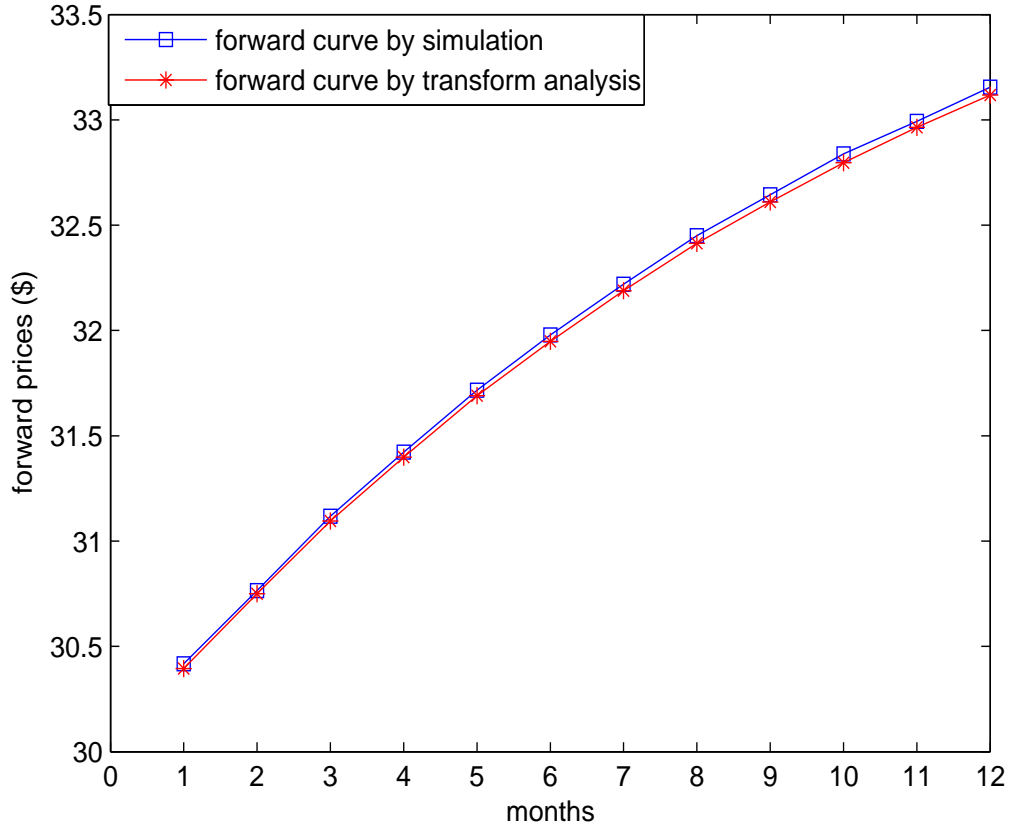
$$F(S_t, t, T) = E^Q[S_T | \mathcal{F}_t] \quad (34)$$

Rewrite (34) as

$$\begin{aligned} F(S_t, t, T) &= E^Q[S_T | \mathcal{F}_t] \\ &= e^{r(T-t)} E^Q[e^{-r(T-t)} S_T | \mathcal{F}_t] \end{aligned} \quad (35)$$

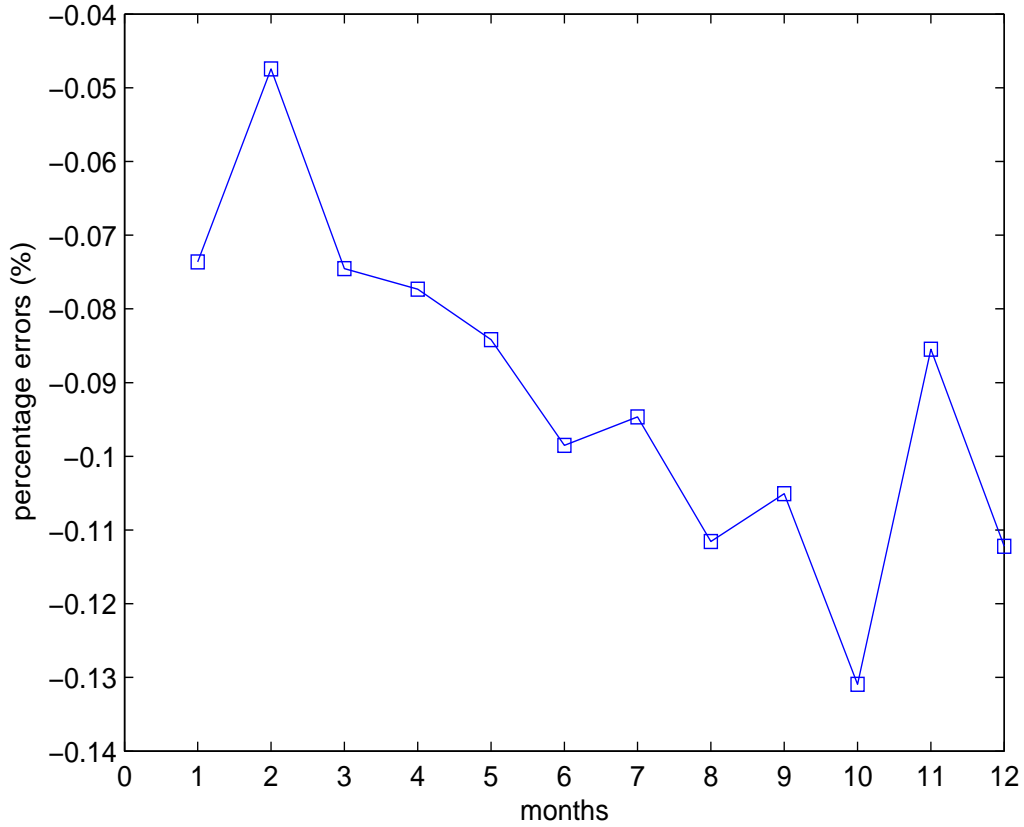
we then have the forward price in terms of the generalized transform function as

$$F(S_t, t, T) = e^{r(T-t)} \Psi(1, Y_t, t, T) \quad (36)$$



**Figure 3:** Forward curves (Contango).  $\mu = 3.4657$ ,  $\alpha = 0.7$ ,  $\sigma = 0.5$ ,  $\nu = 0.03$  and  $Y_0 = 3.4012$

Figure 3 plots the contango forward curves obtained from the simulated data and by the transform analysis. One can see that the forward curve by the transform analysis is pretty close to the simulated forward curve. That is, the transform analysis yields very



**Figure 4:** Percentage errors of forward prices.  $\mu = 3.4657$ ,  $\alpha = 0.7$ ,  $\sigma = 0.5$ ,  $\nu = 0.03$  and  $Y_0 = 3.4012$

good forward prices. Furthermore, I plot the percentage errors in figure 4. The negative values in the figure mean that forward prices obtained by the transform analysis are lower than the simulated forward prices. The maximum percentage error in the figure is only about 0.13%. The main resource of errors is the numerical approximation of integrals in the transform function.

#### 2.4.2.2 Call/Put Option

A “plain vanilla” European call/put option on electricity offers its purchaser the right, but not the obligation, to buy/sell a fixed amount of underlying electricity at a pre-specified strike price  $K$  by the option expiration time. It has a similar payoff structure as a regular call/put option on financial securities or other commodities. The payoff of an electricity

call option is:

$$\text{Payoff of an Electricity Call Option} = \max(S_T - K, 0) \quad (37)$$

at maturity time  $T$ . Thus, the price of the call option at time  $t$  is given by:

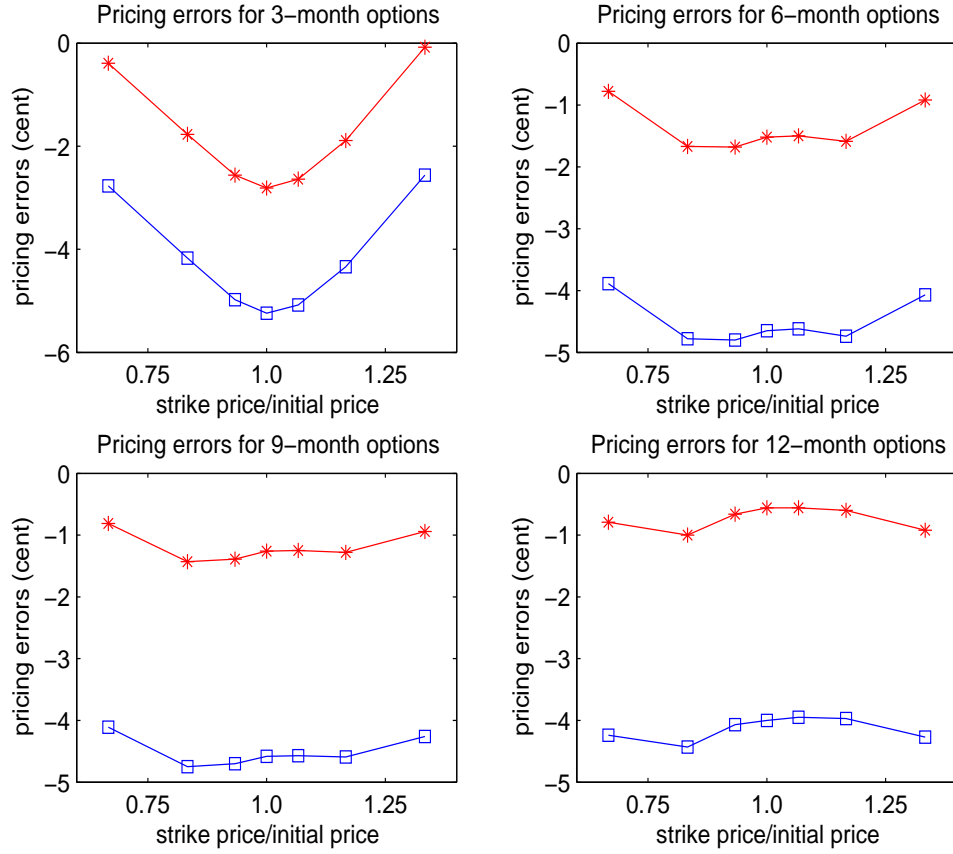
$$\begin{aligned} C(S_t, K, t, T) &= E^Q[e^{-r(T-t)} \max(S_T - K, 0) | \mathcal{F}_t] \\ &= E^Q[e^{-r(T-t)} e^{Y_T} \mathbf{1}_{Y_T \geq \log(K)} | \mathcal{F}_t] - K * E^Q[e^{-r(T-t)} \mathbf{1}_{Y_T \geq \log(K)} | \mathcal{F}_t] \quad (38) \\ &= G_{1,-1}(-\log(K); Y_0, T, \theta) - K * G_{0,-1}(-\log(K); Y_0, T, \theta) \end{aligned}$$

where  $G_{a,b}(x; Y_0, T, \theta)$  is the function given in (32). I plot the pricing errors for a couple of call and put options with different strike prices and time to maturity in figure 5 and figure 6. One can see that the absolute errors are on the level of some cents which again is from the numerical approximation of integrals. The percentage errors are quite small for deep in-the-money options, around 1% for at-the-money options and more than 2% for deep out-of-the-money options. The reason why the percentage errors for deep out-of-the-money options are relatively large is that compared to the absolute errors, the values of deep out-of-the-money options are very small. Moreover, percentage errors decrease as time to maturity increases because with other fixed inputs, options with longer time to maturity have larger values.

In a conclusion, I think that the transform analysis gives very good derivatives prices for the underlying price process being a mean-reverting variance gamma process. Better results are expected if the numerical approximation of integrals in the transform function can be improved.

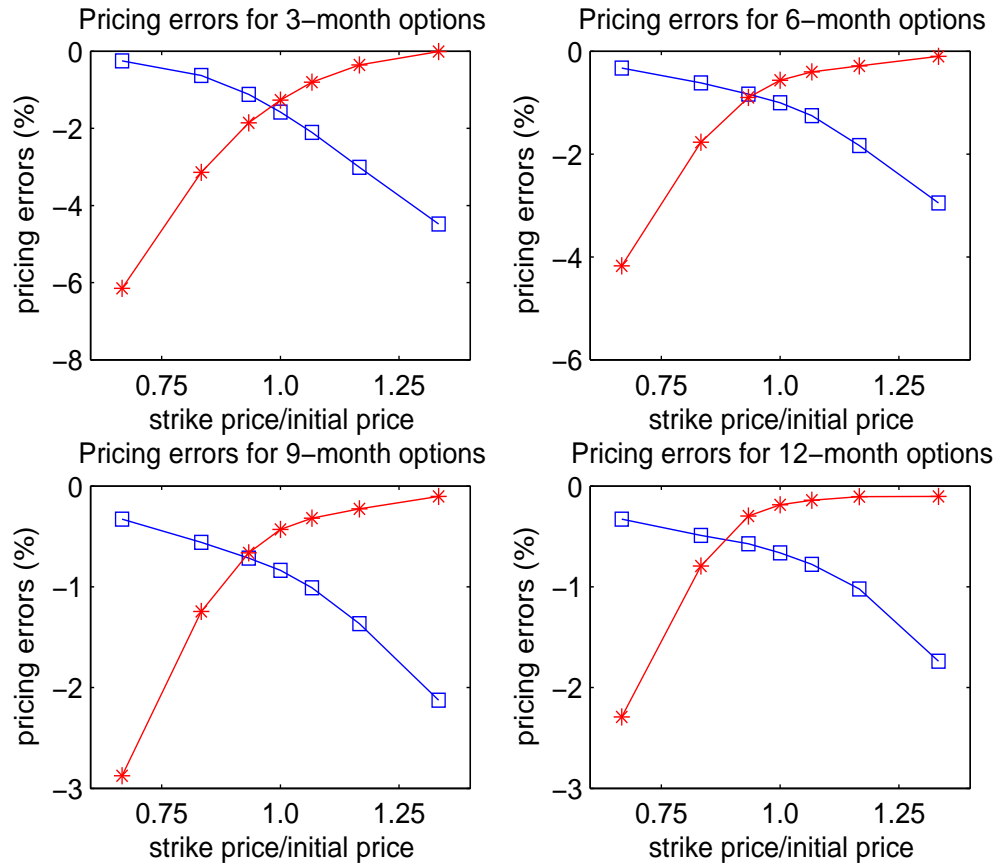
## 2.5 Conclusion

In this chapter, I propose a pure jump process (MRVG) for modeling electricity spot prices. The MRVG process is an extension of a variance gamma process and is expected to be able to capture the salient features of electricity spot prices such as mean-reversion, jumps and spikes. I derive the conditional density function and moment conditions for this new process and demonstrate how the prices of electricity derivatives can be obtained by means of transform analysis. The generalized method of moments and the hybrid Markov chain



**Figure 5:** Option pricing errors (cent). The line with square markers is for call options and the line with star markers for put options.

Monte Carlo method are proposed to estimate model parameters. However, these two methods do not give good estimates of parameters in the simulation study. Thus, an efficient estimation scheme for the MRVG process is needed. Its application in the real world is also for future research.



**Figure 6:** Option pricing errors (%). The line with square markers is for call options and the line with star markers for put options.

## CHAPTER III

# PRICING AND HEDGING ELECTRICITY SUPPLY CONTRACTS: A CASE WITH TOLLING AGREEMENTS

### 3.1 *Introduction*

With the global trend toward the deregulation of the electricity industry, competitive electricity markets have emerged in many countries. In these new electricity markets, the discounted cash flow method, which was a typical approach in the traditional regulated electricity industry, undervalues electricity supply contracts and generation/transmission assets. An appropriate valuation based on market information and effective risk management methodology is important for market participants and the financial institutions, such as banks.

Motivated by Deng and Oren [28] and Tseng et al. [90], I extend the real options approach for valuing power plants to the valuation of electricity contracts with embedded options. A complex electricity contract, such as a *tolling agreement* (or, *tolling contract*), is more challenging to value than a physical power asset (such as a power plant) since the contract can contain contractual constraints that are both operationally set and artificially designed. I formulate a tolling contract as a collection of multiple tolling options (introduced in section 3.3) with constraints on their exercising.

In this chapter, I propose two simulation methods to value a tolling agreement. First of all, I extend a Monte Carlo simulation approach with value function approximation, which is developed in Carriere [15], Longstaff and Schwartz [63], and Tsitsiklis and Van Roy [91] for pricing American options with one single exercising decision to make, to the tolling agreement valuation problem with multiple exercising decisions and side-constraints under general assumptions on the electricity and fuel price dynamics. The similar study by Neinshausen and Hambly [69] extends the duality ideas for American option pricing to the



valuation of multiple-exercise options. For the second approach, I propose exercise policy functions and make exercising decisions according to the values of the policy functions at each time  $t$ . In particular, both approaches are applicable to the models in Deng and Oren [28] and Tseng and Barz [90] with extensions to a wide range of electricity and fuel price assumptions. It can also be applied to other complex energy contract pricing problems such as those in Thompson [88], Jaillet et al. [56], and Keppo [60].

The rest of the chapter is organized as follows. In section 3.2, I describe what a tolling contract is and what constraints I consider in our model. A stochastic dynamic programming valuation model is developed to value a tolling contract in section 3.3. The least square Monte Carlo method and the adaptive policy estimation method are described in section 3.4 and section 3.5, respectively. In section 3.6, I propose a heuristic delta hedging strategy. Numerical results for a tolling agreement on a hypothetical natural gas fired power plant are reported in section 3.7, while conclusion and future works are given in section 3.8.

### ***3.2 Problem Description***

A tolling contract is one of the most innovative structured transactions that the power industry has embraced. A tolling agreement is similar to a common electricity supply contract signed between a buyer (e.g. a power marketer) and an owner of a power plant (e.g. an independent power producer) but with notable differences. For an upfront premium<sup>1</sup> paid to the plant owner, it gives the buyer the right to either operate the power plant or simply take the output electricity during pre-specified time periods subject to certain constraints. In addition to inherent operational constraints of the underlying power plant, there are often other contractual limitations listed in the contract on how the buyer may control the power plant's operations or take the output electricity. For instance, a tolling contract almost always has a clause on the maximum allowable number of power plant restarts as frequent restarting of a generator increases the maintenance costs borne by the

---

<sup>1</sup>Woo et al. [93] provide a statistical benchmark analysis on the reasonableness of the level of such premium based on historical price data of electricity and fuel.

plant owner.

As a tolling agreement gives its buyer the right to take the electricity output of an underlying power plant, subject to certain contractual constraints, holding a tolling contract is equivalent to owning the underlying plant, but with operational flexibility constrained by additional contractual terms. By noting this analogy, I model a tolling agreement as a series of real options on operating a power plant coupled with a restart limit constraint. I elaborate on the modeling details of the operational constraints and the restart limit constraint involved in a tolling contract in this section.

Tolling agreements are written on fossil-fuel power plants. A fossil-fuel plant converts a generating fuel into electricity at a certain conversion rate known as *heat rate*. In brief, heat rate measures the units of the fuel needed for producing one unit of electricity. The lower/higher is the heat rate, the more/less efficient is the power plant. The heat rate is measured in units of MMBtu/MWh where one MMBtu represents one million British thermal units and one MWh stands for one Megawatt (MW) hour of electric energy. In our model, I assume that there is only one power market and one gas market. The owner or any party who operates the power plant has the right, but not the obligation, to generate electricity (e.g., an owner of a merchant power plant). This right-to-generate is known as an *operational option*, which falls into the category of *real options* (see Dixit and Pindyck [30] for more examples of real options). By exercising the operational option, the plant operator receives the spot price of electricity less the heat rate adjusted input fuel cost by buying fuel and selling electricity in their respective spot markets. The “spread” between the electricity price and the heat rate adjusted fuel cost is called *spark spread*. Absent of operational constraints, a rational power plant operator turns on the plant to generate electricity whenever the spark spread (namely, the payoff of the operational option) is positive and shuts down the plant otherwise. Since a *spark spread call option* pays out the positive part of the price difference between the electricity and the generating fuel (namely, the spark spread), the payoff of a power plant at each time epoch  $t$  can be replicated by that of a properly defined spark spread call option (see Deng et al. [25]). Ignoring both operational and contractual constraints, a tolling agreement is simply equivalent to a strip of

spark spread call options with maturity time spanning through the duration of the contract.

The operational constraints in a tolling agreement are naturally tied to those in operating the underlying power plant. Among all aspects of operating a power plant, I consider three major operational characteristics (Wood and Wollenberg [94] offer a good review on power plant operations). First of all, fixed costs are always incurred whenever a power generator is turned on from its “off” state (termed as *startup costs*). Startup costs are generally time dependent. Sometimes, there are costs associated with the turn-off process of a power plant as well, which are called *shutdown costs*. The startup and shutdown costs are fixed costs borne by the tolling contract holder. Secondly, the tolling contract holder usually cannot get electricity output immediately after starting up a power plant. There is a ramp-up delay period  $D$  for a generating unit to reach a certain operating output level starting from the “off” state. Costs incurred during the ramp-up period are also time dependent. Thirdly, a power plant may be operated at a continuum of output levels. At each output level, the generator has a different heat rate. A power plant is usually more efficient (consuming less fuel per unit of electricity generated) when operating in full capacity than running at a lower output level. Therefore, the heat rate of a power plant is a function of the output level.

On the contractual constraints, I use the maximum restart limit described above as one representative example. While the profit of generating electricity comes from the positive spark spread between generated electricity and the input fuel, it is clear that a power plant would only lose money when the spark spread becomes negative, possibly due to too low an electricity price or too high a fuel cost. In times when the spark spread turns so negative that a temporary shutdown of the power generating unit is justified, the operator has to turn off the unit and restart it later when the profit of generating electricity becomes positive again. However, frequent restarts are detrimental to a generation unit, since a restart reduces the unit’s lifetime and increases the likelihood of a forced outage. Due to this fact, there is usually a provision specifying the maximum number of restarts allowed in a tolling contract. Sometimes this constraint is implemented by imposing an extremely high penalty charge on each restart beyond a certain threshold on the cumulative number

of restarts in the contract, effectively capping the total number of restarts at the threshold level. As a result, a tolling contract holder cannot order to shut down the plant at will whenever the electricity spot price is lower than the heat rate adjusted generating fuel cost. Consequently, the value of a tolling agreement is affected by such a constraint.

Intuitively, the value of a tolling contract at any time depends on the *state* of the underlying power plant. The operational characteristics of a power plant provide natural guidelines for defining the state of the plant. The *state* of a tolling agreement encompasses both the operational state of the underlying plant and the operational status related to the contractual obligations. I elaborate on the definition of a power plant's state in a tolling agreement through this example: Suppose a power plant has two output levels: the minimum level and the maximum level, and it takes two phases to ramp up the production from the "off" state to the minimum output level. Then the power plant has five operational states: "off", "ramp-up phase-1", "ramp-up phase-2", "operating at the minimum output level" and "operating at the maximum output level". Consider a tolling agreement on this facility that allows the buyer to restart up to  $n$  times. In this example, the state of the contract at time  $t$  consists of the operational state and the number of allowable restarts left by  $t$ . Figure 7 illustrates all possible states with each circle representing one state and all feasible transitions between any two states of the contract, subject to the number of restarts constraint. Each row in figure 7 corresponds to an operational state of the plant, while every column is tied to the allowable number of restarts left. For instance, the circle at the intersection of the second row and the second column represents a state in which the plant is in the first phase of ramping up and there are  $n - 1$  allowable restarts remaining. With all possible states of a tolling contract defined, I proceed with the problem formulation for valuing a tolling agreement.

### ***3.3 A Stochastic Dynamic Programming Valuation Model***

Consider a tolling contract written on an underlying power plant that has the three operational characteristics discussed in section 3.2. The contract allows no more than  $N$  re-starts of the power plant during its duration of  $T$ . Suppose the contract holder makes decisions

**Figure 7:** State Transition Diagram of a Tolling Contract with the Restart Constraint

on whether to take the output electricity at  $M$  discrete time points  $t_1, t_2, \dots, t_M$  over the horizon  $[0, T]$  where  $0 = t_1 < t_2 < \dots < t_M = T$ .  $N$  is very small compared to  $M$ . The fact that the holder may take the electricity at any time  $t$  is modeled by letting  $M$  be an arbitrarily large integer. Since the electricity and the fuel are traded in the open markets, the holder elects to take the electricity whenever the spark spread is positive due to the no-arbitrage principle. As a result, the optimal take-or-not (and, quantity-to-take) decisions by the contract holder correspond exactly to the optimal produce-or-not (and, quantity-to-produce) decisions by the plant operator under the objective of maximizing the cumulative profit of the power plant subject to the tolling contract provisions.

Let us define a *tolling option* in a tolling contract to be the right of a contract holder to start taking the output electricity of the underlying plant at any time with self-supplied generating fuel, and the obligation, after exercising the right-to-take, to continuously take electricity (possibly in varying quantities) until she/he chooses to stop. The value of a tolling

contract is therefore equal to the maximized total payoff associated with all exercised tolling options subject to a constraint that no more than  $N$  tolling options can be exercised during the life of the tolling contract. In exercising a tolling option, two sequential decisions need to be made: the first being when to start taking electricity, and the second being when to stop. By no-arbitrage, the underlying plant is started (re-started) at the beginning of an exercised tolling option and shut down at the termination time. The contract holder is responsible for the corresponding startup and shutdown costs. If there were no startup or shutdown costs or other operational and contractual constraints of a power plant, then the optimal decisions in exercising a tolling option would be to start taking electricity whenever the spark spread turns positive and to stop doing so whenever the spark spread turns negative. In such an ideal case, a tolling option is simply a series of spark spread call options with the longest maturity time given by the first time of hitting zero by the spark spread with a positive initial value.

As explained in Dixit and Pindyck [30], the real options (in this case, tolling options) can be valued by a stochastic dynamic programming (SDP) approach. When the payoff of the real options is perfectly replicated by traded financial instruments such as forward contracts on electricity and fuel, the correct discount rate used in the SDP approach needs to be the risk-free interest rate and the SDP approach becomes equivalent to the contingent claim analysis for option-pricing as developed in Black and Scholes [6], Merton [70], and Harrison et al. [50]. In the case where the available traded financial instruments cannot achieve perfect hedging (i.e., incomplete market), the discount rate is obtained by adding a risk premium to the risk-free rate.

### 3.3.1 Formulation of the Tolling Contract Valuation

The following notations are used throughout the paper.

- $V_t$  : value of the tolling contract at time  $t$ ,
- $a_t$  : operational action taken at time  $t$  by an expected profit-maximizing power plant operator,
- $R_t$  : payoff of the operational option at time  $t$ ,
- $n_t$  : number of power plant re-starts left at time  $t$ ,
- $w_t$  : operational state of the power plant at time  $t$ ,
- $\Theta_t^P$  : state of the tolling contract,  $(w_t, n_t)$ , at time  $t$ ,

- $X_t, Y_t$  : natural logarithm of electricity and the fuel prices at time  $t$ , respectively,
- $\Theta_t^S$  : log-price vector  $(X_t, Y_t)$ ,
- $\Theta_t$  : vector  $(X_t, Y_t, w_t, n_t)$  representing the state of the world.

#### 3.3.1.1 Value Function

From here on, I refer to  $X_t$  and  $Y_t$  as prices with the understanding that they represent log-prices. Suppose that the price vector  $(X_t, Y_t) \in \mathbb{R}^2$  evolves according to a Markov process defined in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with initial value  $(X_0, Y_0)$  at time 0. The  $\sigma$ -field generated by the stochastic process  $\{(X_s, Y_s) : 0 \leq s \leq t\}$ , denoted by  $\mathcal{F}_t \subset \mathcal{F}$ , forms a filtration  $\mathbb{F}$  over time interval  $[0, T]$ . Let  $A_t$  denote the set of admissible operations available to the plant operator at time  $t$  and  $R_t$  be the payoff of the time- $t$  operational option. Based on the previous analysis, the value of a tolling contract at time  $t \in \{t_1, t_2, \dots, t_M\}$  is given by

$$V_t = \max_{\{a_{t_i} \in A_{t_i}, \dots, a_{t_M} \in A_{t_M}\}} E\left[\sum_{I=t}^T e^{-r(t_I-t)} R_{t_I} | \mathcal{F}_t\right] \quad (39)$$

where  $t = t_i$  and  $r$  is a discount rate.  $R_t$  can be interpreted as the operating profit during  $t_i$  to  $t_{i+1}$ . It depends on both the operational action and the state of the world, namely,  $R_t \equiv R(a_t, \Theta_t) \equiv R(a_t, X_t, Y_t, w_t, n_t)$ . While the range of  $(X_t, Y_t)$  is  $\mathbb{R}^2$  and  $n_t \in W_N \equiv \{0, 1, 2, \dots, N\}$ , I need to introduce  $w_t$  and  $a_t$  before defining  $R(a_t, \Theta_t)$ .

For the ease of exposition, I make further simplifying assumptions on the operating characteristics which can be readily generalized. Specifically, the underlying power plant

has one “off” state and two output states: the minimum output state of generating  $\underline{Q}$  MW per time unit (with heat rate  $\underline{Hr}$ ) and the maximum output state of generating  $\overline{Q}$  MW per time unit (with heat rate  $\overline{Hr}$ ). Since a power plant works more efficiently at the maximum output level than at the minimum output level,  $\overline{Hr}$  is smaller than  $\underline{Hr}$ . I also assume no-delay and no-cost in switching between the minimum and the maximum output levels. A start-up cost  $c_{up}$  is incurred whenever the power plant is turned on from the “off” state. Recall there is a delay (or ramp-up) period of  $D$  (called the ramp-up time) before a power plant can output electricity after the plant is turned on from the “off” state. Without loss of generality, I assume that  $D$  is a multiple of  $\Delta t$  where  $\Delta t$  is the length of the small time intervals over which the operating decisions are made. Let  $K_D$  denote  $\frac{D}{\Delta t}$ . Then there are  $K_D - 1$  ramp-up states. During the ramp-up period, the cost is  $c_r(Y_t)$  per time unit at time  $t$ , which is a positive increasing function of the generating fuel price  $Y_t$ . A shut-down cost  $c_{down}$  is incurred whenever the power plant is turned off. To summarize, I use a set  $W_D \equiv \{0, 1, \dots, K_D - 1, K_D\}$  to represent the  $(K_D + 1)$  possible operational states. Specifically,  $w_t$  takes on  $(K_D + 1)$  possible values at time  $t$ .

- $w_t = 0$  : The power plant is in *off* state at time  $t$ .
- $w_t = i$  : The power plant is *on* but in the  $i^{th}$  stage of the *ramp-up* period  $D$  at time  $t$  for  $i \in \{1, 2, \dots, K_D - 1\}$ .
- $w_t = K_D$  : The power plant is *on and ready* to generate electricity outputs at time  $t$ .

The operational action of the plant operator,  $a_t$ , has three possible choices  $a^i$  ( $i = I, II, III$ ) corresponding to the operational states. The admissible action set  $A_t$  in (39) is a subset of  $A \equiv \{a^I, a^{II}, a^{III}\}$  for all time  $t$ .

- $a^I$ : The operator operates the power plant at the maximum capacity level. The plant generates  $\overline{Q} \cdot \Delta t$  units of electricity in time  $\Delta t$  with an operating heat rate of  $\overline{Hr}$  if it is not in a ramp-up stage; otherwise it generates 0 units of electricity.
- $a^{II}$ : The operator keeps the power plant running at the minimum capacity level. The plant generates  $\underline{Q} \cdot \Delta t$  units of electricity in time  $\Delta t$  with an operating heat rate of



$\underline{Hr}$  if it is not in a ramp-up stage; otherwise it generates 0 units of electricity.

- $a^{III}$ : The operator turns the power plant off from any non-“off” state or keeps it off if it is in the off state.

The operating profit of the power plant at any time  $t$ ,  $R(a, x, y, w, n) : A \times R^2 \times W_D \times W_N \rightarrow R^1$ , is defined as follows. The operational characteristics described above are reflected in the definition of  $R(a, x, y, w, n)$ .

- When  $w = 0$ ,

$$R(a, x, y, 0, n) = \begin{cases} -\infty, & \text{if } a_t = a^I \text{ or } a^{II}, n = 0, \forall (x, y). \\ -c_{up} - c_r(y) \cdot \Delta t, & \text{if } a = a^I \text{ or } a^{II}, n \geq 1, \forall (x, y). \\ 0, & \text{if } a_t = a^{III}, \forall n, \forall (x, y). \end{cases} \quad (40)$$

- When  $w = 1, 2, \dots, (K_D - 1)$ ,

$$R(a, x, y, w, n) = \begin{cases} -c_r(y) \cdot \Delta t, & \text{if } a = a^I, \forall n, \forall (x, y). \\ -c_r(y) \cdot \Delta t, & \text{if } a = a^{II}, \forall n, \forall (x, y). \\ -c_{down}, & \text{if } a = a^{III}, \forall n, \forall (x, y). \end{cases} \quad (41)$$

- When  $w = K_D$ ,

$$R(a, x, y, K_D, n) = \begin{cases} \overline{Q} \cdot \Delta t \cdot [e^x - \overline{Hr} \cdot e^y], & \text{if } a = a^I, \forall n, \forall (x, y). \\ \underline{Q} \cdot \Delta t \cdot [e^x - \underline{Hr} \cdot e^y], & \text{if } a = a^{II}, \forall n, \forall (x, y). \\ -c_{down}, & \text{if } a = a^{III}, \forall n, \forall (x, y). \end{cases} \quad (42)$$

### 3.3.1.2 Hamilton-Jacobi-Bellman Equations for the Value Function

With  $\{(X_t, Y_t) : t \geq 0\}$  being a Markov process and  $R_t$  defined by (40)-(42), the contract value function  $V_t$  in (39) simplifies to a function of the state variables  $(X_t, Y_t, w_t, n_t)$  at time  $t$ , namely,  $V_t = V_t(X_t, Y_t, w_t, n_t) \equiv V_t(\Theta_{t+1}^S, \Theta_{t+1}^P)$ . Moreover, the value function  $V_t(X_t, Y_t, w_t, n_t)$  satisfies the following Hamilton-Jacobi-Bellman equations in all possible states of the contract  $\Theta_t^P$  at every  $t_i \in \{t_1, t_2, \dots, t_{M-1}\}$ . Let  $E_t[\cdot]$  denote the conditional expectation operator  $E[\cdot | \mathcal{F}_t]$ .

- For every price vector  $\Theta_t^S = (X_t, Y_t) \in R^2$  and the state of the contract  $\Theta_t^P$  being  $(w_t, n_t) = (0, n_t)$  for every  $n_t \geq 1$  at  $t = t_i$ ,

$$V_t(\Theta_t^S, \Theta_t^P) = \max_{a_t \in A_t} \begin{cases} a_t = a^I : -c_{up} - c_r(Y_t)\Delta t + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{II} : -c_{up} - c_r(Y_t)\Delta t + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{III} : e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_t^P)] \end{cases} \quad (43)$$

where  $\Delta t = (t_{i+1} - t_i)$  and  $\Theta_{t_{i+1}}^P = (1, n_t - 1)$ .

- For every price vector  $\Theta_t^S = (X_t, Y_t) \in R^2$  and every state of the contract  $\Theta_t^P = (w_t, n_t) \in \{(W_D \setminus \{0, K_D\}) \times W_N\}$  at  $t = t_i$ ,

$$V_t(\Theta_t^S, \Theta_t^P) = \max_{a_t \in A_t} \begin{cases} a_t = a^I : -c_r(Y_t)\Delta t + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{II} : -c_r(Y_t)\Delta t + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{III} : -c_{down} + e^{-r\Delta t} E_t[V_{t_{i+1}}(X_{t_{i+1}}, Y_{t_{i+1}}, 0, n_t)] \end{cases} \quad (44)$$

where  $\Delta t = (t_{i+1} - t_i)$  and  $\Theta_{t_{i+1}}^P = (w_t + 1, n_t)$ .

- For every price vector  $\Theta_t^S = (X_t, Y_t) \in R^2$  and the state of the contract  $\Theta_t^P = (K_D, n_t)$  for every  $n_t \in W_N$  at  $t = t_i$ ,

$$V_t(\Theta_t^S, \Theta_t^P) = \max_{a_t \in A_t} \begin{cases} a_t = a^I : \overline{Q}\Delta t[e^{X_t} - \overline{Hr} \cdot e^{Y_t}] + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{II} : \underline{Q}\Delta t[e^{X_t} - \underline{Hr} \cdot e^{Y_t}] + e^{-r\Delta t} E_t[V_{t_{i+1}}(\Theta_{t_{i+1}}^S, \Theta_{t_{i+1}}^P)] \\ a_t = a^{III} : -c_{down} + e^{-r\Delta t} E_t[V_{t_{i+1}}(X_{t_{i+1}}, Y_{t_{i+1}}, 0, n_t)] \end{cases} \quad (45)$$

where  $\Delta t = (t_{i+1} - t_i)$  and  $\Theta_{t_{i+1}}^P = (K_D, n_t)$ .

Since the contract has no value to the holder in two scenarios: a) when the contract reaches its expiration time  $T$ , regardless of whether the power plant is *off* or not; and, b) when the power plant is *off* and the number of allowable re-starts is 0, the boundary conditions are

$$\begin{aligned} V_t(X_t, Y_t, 0, 0) &\equiv 0, \quad \forall (X_t, Y_t) \in R^2 \text{ and } \forall t \in \{t_1, t_2, \dots, t_M\}, \\ V_T(X_T, Y_T, w_T, n_T) &\equiv 0, \quad \forall (X_T, Y_T) \in R^2 \text{ and } \forall (w_T, n_T) \in W_D \times W_N, \\ V_t(X_t, Y_t, w_t, -1) &\equiv -\infty, \quad \forall (X_t, Y_t) \in R^2 \text{ and } \forall w_t \in W_D. \end{aligned} \quad (46)$$

### 3.3.2 Specification of the Underlying Commodity Price Processes

In the SDP formulation, price processes  $\{(X_t, Y_t) : t \geq 0\}$  are key components. To illustrate the impacts of the commodity price assumptions on the contract valuation, I specify two alternative underlying commodity price models and investigate the effects of different price models on the valuation of a tolling contract. I shall see from the numerical examples in section 3.7 that different commodity price models systematically bias the value of a tolling agreement against one another.

For modeling the electricity price  $X_t$ , I consider both a simple mean-reverting process and a mean-reverting jump-diffusion process which are adapted to the filtration  $\mathbb{F}$ . The generating fuel price  $Y_t$  is modeled by an adapted simple mean-reverting process.

*Case 1: both  $\{X_t : t \geq 0\}$  and  $\{Y_t : t \geq 0\}$  are mean-reverting processes.*

$$\begin{aligned} dX(t) &= \alpha_1(\mu_1 - X(t))dt + \sigma_1 dW_1(t) \\ dY(t) &= \alpha_2(\mu_2 - Y(t))dt + \rho\sigma_2 dW_1(t) + \sqrt{1 - \rho^2}\sigma_2 dW_2(t) \end{aligned} \quad (47)$$

where  $\alpha_i$  ( $i = 1, 2$ ) are mean-reverting speeds,  $\mu_i$  ( $i = 1, 2$ ) are long-term means of log-prices,  $\sigma_i$  ( $i = 1, 2$ ) are price volatilities,  $\rho$  is the instantaneous correlation coefficient between the two price processes, and  $W_i(t)$  ( $i = 1, 2$ ) are independent standard Brownian motions. Regarding parameter estimation, I observe that, over a small time increment  $\Delta t$ , the diffusion terms  $\sigma_i dW_i(t)$  ( $i = 1, 2$ ) of  $dX(t)$  and  $dY(t)$  in (47) are simply Normal random variables with mean zero and variance  $\sigma_i^2 \Delta t$  ( $i = 1, 2$ ). Thus, for price  $X(t)$  sampled at time  $0, \Delta t, 2\Delta t, \dots, n\Delta t$ , I can regress the log return  $\Delta X \equiv X(t + \Delta t) - X(t)$  over  $\Delta t$  onto  $X(t)$  scaled by  $\Delta t$  to estimate  $\mu_1$  and  $\alpha_1$ . The standard error of the residuals provides an estimate of  $\sigma_1$  (See Clewlow and Strickland [18]). The parameters of  $Y(t)$  can be estimated in the same fashion.

*Case 2:  $X(t)$  is a mean-reverting jump-diffusion process and  $Y(t)$  is a mean-reverting process:*

$$\begin{aligned} dX(t) &= \alpha_1(\mu_1 - X(t))dt + \sigma_1 dW_1(t) + \kappa dq(t) \\ dY(t) &= \alpha_2(\mu_2 - Y(t))dt + \rho\sigma_2 dW_1(t) + \sqrt{1 - \rho^2}\sigma_2 dW_2(t) \end{aligned} \quad (48)$$

where  $\alpha_i$ ,  $\mu_i$ ,  $\sigma_i$ ,  $\rho$  and  $W_i(t)$  have the same interpretations as those in (47),  $q(t)$  is a Poisson process with intensity  $\phi$  independent of everything else, and  $\kappa$  denotes a Normal random

variable  $N(\bar{\kappa}, \gamma)$ .  $q(t)$  and  $\kappa$  model the occurrence and size of jumps, respectively. To estimate the parameters for  $X(t)$  in (48), I adopt a heuristic method outlined in Clewlow and Strickland [18]. The first step is to determine a reasonable threshold level of return beyond which a log-return is considered to be due to a price jump. For instance, in the numerical examples in section 3.7, if a log-return is at least three times of the standard deviation away from the mean of the log-returns, I determine that it results from a price jump. With this threshold, I examine all return data and take out those associated with jumps from the entire data set. I then repeat the previous step with a new threshold based on the mean of the remaining data until no more jump returns can be removed. The mean-reverting and diffusion parameters  $\alpha_1, \mu_1$  and  $\sigma_1$  are estimated from the resulting data set by applying the same procedure as the one used for the simple mean-reverting model (47). I next count the total number of jumps that I removed from the original data set and divide that number by the length of the sample period to obtain the jump frequency  $\phi$ . The mean and the standard deviation of the jump size  $\kappa$  are also computed from these jumps.

### ***3.4 Least Square Monte Carlo Method***

For pricing a financial option, there are two classical numerical approaches: pricing by a lattice (See Cox et al. [19]) and pricing by Monte Carlo simulation (See Boyle [9]). As the valuation problem is framed as pricing a series of real options, both the lattice method and the simulation method can be applied in theory. However, when the underlying price process has a jump component, as in the case of (48), it is difficult to construct an end-recombining lattice which converges to the continuous-time model in distribution. Moreover, if the valuation horizon is long, the computational effort of the lattice approach can be huge, thus making the computational time prohibitively long. As an alternative, I propose a Monte Carlo simulation-based method using value function approximation by a least-squares regression technique, which is developed in Carriere [15], Longstaff and Schwartz [63], and Tsitsiklis and Van Roy [91] for pricing American-style options, to solve the dynamic programming problem of tolling contract valuation.

### 3.4.1 Commodity Price Process Simulation

To simulate the electricity and the generating fuel prices, I apply the Euler method in Higham [53] to discretize the continuous-time models defined in (47) and (48). The two correlated price processes in the two cases are simulated through the stochastic difference equations (49) and (50).

*Case 1:  $X(t)$  and  $Y(t)$  are mean-reverting processes:*

$$X(t + \Delta t) = X(t) + \alpha_1(\mu_1 - X(t))\Delta t + \sigma_1\epsilon_1\sqrt{\Delta t} \quad (49)$$

$$Y(t + \Delta t) = Y(t) + \alpha_2(\mu_2 - Y(t))\Delta t + \sigma_2\rho\epsilon_1\sqrt{\Delta t} + \sqrt{1 - \rho^2}\sigma_2\epsilon_2\sqrt{\Delta t}$$

*Case 2:  $X(t)$  is an mean-reverting jump-diffusion process and  $Y(t)$  is an mean-reverting process:*

$$X(t + \Delta t) = X(t) + \alpha_1(\mu_1 - X(t))\Delta t + \sigma_1\epsilon_1\sqrt{\Delta t} + I_{(U < \phi\Delta t)}(\bar{\kappa} + \gamma\epsilon_3) \quad (50)$$

$$Y(t + \Delta t) = Y(t) + \alpha_2(\mu_2 - Y(t))\Delta t + \sigma_2\rho\epsilon_1\sqrt{\Delta t} + \sqrt{1 - \rho^2}\sigma_2\epsilon_2\sqrt{\Delta t}$$

where  $\epsilon_1$ ,  $\epsilon_2$  and  $\epsilon_3$  are independent standard Normal random variables, and  $U$  is a uniform random variable in  $[0, 1]$ .

The hourly power prices from 06:00 to 22:00 each day are known as on-peak prices, and those for the remaining hours in a day are known as off-peak prices. One important fact about power prices is the time-of-day effect: on-peak prices are much higher than off-peak prices. To reflect this fact, I discretize the time horizon into alternating small intervals of two different lengths. One length is 16-hour, corresponding to the peak-hour interval in one day and the other is 8-hour which corresponds to the off-peak interval. Each peak-hour interval is followed by an off-peak interval and vice versa. Finally, I simulate the “daily” prices over these small time intervals according to (49) and (50). The on-peak and off-peak prices are obtained by multiplying the daily prices over on-peak and off-peak intervals with different scaling factors  $k_{on}$  and  $k_{off}$ , respectively. In section 3.7, based on historical data, I choose  $k_{on} = 1.2$  for the on-peak price and  $k_{off} = 0.6$  for the off-peak price, so the average price over 24 hours in one day equals the daily price simulated according to (49) and (50).

### 3.4.2 Forward Curve Modeling

In practice, the value of a tolling agreement depends heavily on market forward curves and volatility term structure. To make our real option model applicable to the real world, I make use of market forward information of prices and volatilities at time 0 when simulating the underlying prices.

Implied volatility refers to the volatility parameter corresponding to a given call option value through the Black-Scholes call option pricing formula. It is obtained by inverting the Black-Scholes pricing formula using the given call option price and other known parameters as inputs. Let  $C_t$  denote the at-the-money call option price observed in the market at time 0 that will expire in month  $t$ . Then the implied volatility  $\sigma_{0,t}$  can be obtained by inverting the Black-Scholes pricing formula using the call option price  $C_t$  and other known parameters as inputs. Thus forward volatilities  $\sigma_{t-1,t}$  for  $t = 2, \dots, T$  where  $T$  is the maturity time of a tolling agreement can be obtained by solving the following approximate equation for  $t = 2, \dots, T$ :

$$\sigma_{0,t}^2 * t = \sigma_{0,t-1}^2 * (t-1) + \sigma_{t-1,t}^2 \quad (51)$$

Let  $F_{0,t}, t = 1, \dots, T$  denote the forward curve observed at time 0. When simulating spot prices of underlying assets for month  $t$  according to (49) or (50), I use  $\sigma_{t-1,t}$  and  $F_{0,t}$  as inputs for  $\sigma$  and  $\mu$  in (49) and (50). After simulation, I further adjust the simulated data for month  $t$  by adding the spread

$$F_{0,t} - \text{the mean of simulated spot prices for month } t \quad (52)$$

Since the forward curve usually exhibits seasonality, the adjusted spot prices will have the feature of seasonality as well.

### 3.4.3 Value Function Approximation by Least-squares Regression

In solving  $V_t$  through (43)-(46) in either of the two cases, it is a difficult task to compute the exact conditional expectation of the value function. To overcome this difficulty, I propose to approximate the conditional expectation of the value function in our problem by expanding it with respect to a set of complete basis functions and obtaining the expansion coefficients

through least-squares regression. Such an approach is developed in Carriere [15], Longstaff and Schwartz [63], and Tsitsiklis and Van Roy [91] for pricing American options with one single exercising decision to make. I show that it is applicable to the tolling agreement valuation problem, which also involves multiple exercising decisions and a constraint on the total number of allowable exercises. The key is to use a truncated expansion of the conditional expected value function with respect to a set of basis functions for the approximations in (43)-(46). As the approximation is a linear combination of the basis functions, a least-squares regression can be applied to obtain the expansion coefficients. Specifically, I regress a conditional expected value function onto basis functions which are functions of the underlying price variables. Polynomial functions up to the third order are chosen as the basis functions for the implementation specified in (53).

Let  $\tilde{V}_t(X_t, Y_t, w, n)$  denote  $e^{-r\Delta t}E_t[V_{t+1}(X_{t+1}, Y_{t+1}, w, n)]$ . Given a set of simulated sample paths  $\Lambda \equiv \{(X_t(\omega_i), Y_t(\omega_i)) : \omega_i \in \Omega (i = 1, 2, \dots, J), 0 \leq t \leq T\}$ , I regress the vector consisting of  $e^{-r\Delta t}V_{t+1}(X_{t+1}(\omega_i), Y_{t+1}(\omega_i), w, n)$  ( $i = 1, 2, \dots, J$ ) onto a set of polynomial functions of  $(X_t(\omega_i), Y_t(\omega_i))$  as given in the right hand side (RHS) of (53) and obtain the coefficients  $\{b_i^t(w, n) : i = 0, 1, \dots, 9; (w, n) \in W_D \times W_N\}$  by a least-squared regression.

$$\begin{aligned}
e^{-r\Delta t}E_t[V_{t+1}(X_{t+1}, Y_{t+1}, w, n)] &\cong b_0^t(w, n) + b_1^t(w, n)e^{X_t} + b_2^t(w, n)e^{Y_t} \\
&\quad + b_3^t(w, n)e^{2X_t} + b_4^t(w, n)e^{2Y_t} + b_5^t(w, n)e^{X_t}e^{Y_t} \\
&\quad + b_6^t(w, n)e^{3X_t} + b_7^t(w, n)e^{3Y_t} + b_8^t(w, n)e^{2X_t}e^{Y_t} \\
&\quad + b_9^t(w, n)e^{X_t}e^{2Y_t} \quad \forall (w, n) \in W_D \times W_N
\end{aligned} \tag{53}$$

I then use the RHS of (53) to approximate  $\tilde{V}_t(X_t, Y_t, w, n)$  and substitute it into the term  $e^{-r\Delta t}E_t[V_{t+1}(\Theta_{t+1}^S, \Theta_{t+1}^P)]$  in (43)-(45) to solve for  $V_t(X_t, Y_t, w_t, n_t)$  by backwards induction starting with terminal condition (46). Working backwards until reaching the starting time, I obtain the initial value of a tolling agreement.

### 3.5 Adaptive Policy Estimation Method

The adaptive policy estimation method also is a simulation-based approach, which was proposed to price American-style options in Deng and Lee [26]. As I discuss in section 3.2, a tolling contract can be formulated as a collection of multiple tolling options with constraints in their exercising. The core of the adaptive policy estimation method is that I parameterize some exercise policy functions, which are functions of time and underlying prices. Setting the policy functions to be zero forms some boundaries separating the state variable space into on-region, off-region and no-action-region. One determines when to turn on (turn off) the underlying power plant (exercising the tolling options) according to the values of policy functions. Given current state variables, if the value of turn-on policy function is greater than zero, the power plant should be turned on. Similarly, if the value of turn-off policy function is smaller than zero, one should turn off the power plant. Since I assume that there is no cost and no delay on switching the output level, operating the power plant under different output levels do not affect exercise decisions, only the payoff at any time  $t$ .

#### 3.5.1 Policy Function

An operator who seeks to maximize the operating profit usually turns on the plant when the spark spread is positive and shuts down the plant when the spark spread is negative, subject to some constraints. Therefore, I construct our policy functions based on the spark spread  $X(t) - Hr * Y(t)$ . I propose policy functions  $f_i(t; X_t, Y_t, n_t) : [0, T] \times R^2 \times W_N \rightarrow R, i \in \{on, off\}$  linear on the price vector  $\Theta_t^S = (X_t, Y_t) \in R^2$  for turn-on policy and turn-off policy respectively as follows:

$$f_i(t; X_t, Y_t, n_t) = K_i(t, n_t)X_t + C_i(t, n_t) - Y_t, \quad i \in \{on, off\} \quad (54)$$

where the slope  $K_i(t, n_t)$  and the interception  $C_i(t, n_t)$  both are expected to be functions of time  $t$  for each  $n_t \in W_N$ . To reflect the fact that on-peak electricity prices are much higher than off-peak prices, I choose the following piecewise linear functions of time  $t$  to model  $K_i(t, n_t)$  and  $C_i(t, nt)$  for each  $n_t \in W_N$ :



$$K_{j,on}(t) = \begin{cases} k_{j,on}^1 t + k_{j,on}^2, & t \leq \tau_{j,on}, \quad j \in \{\text{on-peak}, \text{off-peak}\} \\ k_{j,on}^1 \tau_{j,on} + k_{j,on}^2, & t > \tau_{j,on}, \quad j \in \{\text{on-peak}, \text{off-peak}\} \end{cases} \quad (55)$$

$$C_{j,on}(t) = \begin{cases} c_{j,on}^1 t + c_{j,on}^2, & t \leq s_{j,on}, \quad j \in \{\text{on-peak}, \text{off-peak}\} \\ c_{j,on}^1 s_{j,on} + c_{j,on}^2, & t > s_{j,on}, \quad j \in \{\text{on-peak}, \text{off-peak}\} \end{cases} \quad (56)$$

$$K_{j,off}(t) = \begin{cases} k_{j,off}^1 t + k_{j,off}^2, & t \leq \tau_{j,off}, \quad j \in \{\text{on-peak}, \text{off-peak}\} \\ k_{j,off}^1 \tau_{j,off} + k_{j,off}^2, & t > \tau_{j,off}, \quad j = \text{on-peak} \\ k_{j,off}^3 t + k_{j,off}^4, & t > \tau_{j,off}, \quad j = \text{off-peak} \end{cases} \quad (57)$$

$$C_{off}(t) = \begin{cases} c_{off}^1 t + c_{off}^2, & t \leq s_{off} \\ c_{off}^1 s_{off} + c_{off}^2, & t > s_{off} \end{cases} \quad (58)$$

In (55)-(58), I have 23 parameters in a total to estimate, each of which is a mass function of the number of restarts  $n$ . One thing to notice is that all piecewise linear functions in (55)-(58) have a horizontal second piece except the slope of the off-peak turn-off policy function  $K_{\text{off-peak},off}(t, n_t)$ , which has a nonzero slope for its second piece. Another issue is that I do not distinguish on-peak and off-peak for the interception of the turn-off policy function. The reason I do not employ more general piecewise linear functions in (55)-(58) is that based on our empirical experiments, there is no need to introduce more parameters in our valuation model.

### 3.5.2 The Adaptive Policy Search Algorithm

With the policy functions proposed above, I can value a tolling contract by the following procedure:

1. Pick a sufficient large number  $N$ .  $\Delta t = T/N$ . Simulate  $N_1$  sample paths of  $X_t$  and  $Y_t$  over time horizon  $[0, T]$  according to (49) or (50).
2. Assume that each parameter has a discrete distribution over some interval. Theoretically, any distribution could be used as an initial distribution. For simplicity, I start with a uniform distribution. A reasonable lower bound and upper bound of the

interval for each parameter needs to be identified. Each interval is equally partitioned into  $M$  sub-intervals and the mid-point of each sub-interval is assigned a weight 1 (probability  $1/M$ ).

3. Let  $k = 0$ . Initialize  $P^{\{0\}}$  to be a joint probability mass function of 23 independent discrete uniform random variables.
4. Sample  $N_P$  sets of parameters according to  $P^{\{k\}}$
5. Using parameters obtained above to generate  $N_P$  turn-off and turn-on policies
6. For each sample path, determine the status of the power plant at each time spot  $t$ . Assume that the initial status of the power plant at time zero is  $(w_0, n_0) = (0, n)$ . The power plant will stay in the status  $(0, n)$  until the first positive value of the turn-on policy is found at time  $t_1$ . The status now changes to  $(w(t_1), n(t_1)) = (1, n-1)$ . If the plant is not turned off at time  $t_1+1$ , the status will be  $(w(t_1+1), n(t_1+1)) = (2, n-1)$ . The status will change until time  $t_1 + K_D$  because the plant is in the ramp up period. At time  $t_1 + K_D$ , the plant will be in the status  $(w(t_1 + K_D), n(t_1 + K_D)) = (K_D, n-1)$ , and it will maintain this status until it is turned off. Once the plant is turned on at time  $t_1$ , it will not be turned off until the first negative value of the turn-off policy is found at time  $t_2 > t_1$ . The status of the plant at time  $t_2$  becomes  $(w(t_2), n(t_2)) = (0, n-1)$ . The plant is off again and only  $n-1$  restarts are left. The status  $(0, n-1)$  will keep until the next positive value of the turn-on policy at time  $t_3 > t_2$  arrives. Whenever the plant goes into the “off” state, and  $n_t = 0$ , that is, the status of the plant becomes  $(0, 0)$ , the plant will stay there until the maturity time  $T$ .
7. Apply each exercising boundary to the tolling agreement over the set of  $N_1$  sample paths and obtain an approximate value  $V_{n_b}(n_b = 1, 2, \dots, N_P)$  of the true contract value. Find a “good” set of exercise boundaries based on  $V_{[\rho N_P]}$  where  $V_{[\rho N_P]}$  denotes the upper  $\rho$  quantile of all the approximate option values. Let  $Q_i^{(k)}$  denote the exercising boundary yielding the  $i^{th}$ -highest contract value where  $k$  is the iteration number for policy search.

8. Compute a new PMF  $P^{\{*\}}$  based the good policy sets  $Q_i^{(k)}$  where  $i = 1, \dots, \lceil \rho N_P \rceil$ . For example, let  $x_i, i = 1, \dots, M$  denote the  $M$  sub-intervals (mid-points) for some parameter  $x$ , and each sub-interval has a weight  $u_i^k, i = 1, \dots, M$  so that the total weight is  $U = \sum_{i=1}^M u_i^k$ . If  $x_i$  appears  $p \in \{0, 1, \dots, \lceil \rho N_P \rceil$  times among the good policy sets, add  $p$  to its weight  $u_i$ , where  $i=1, \dots, M$ . Then the total weight is  $U' = U + \lceil \rho N_P \rceil$ . The mass probabilities can be obtained by dividing the updated  $u_i$  by  $U'$ .
9. Exit the loop if the sum of squared errors of expectations of all parameters under  $P^{\{k\}}$  and  $P^{\{*\}}$  is smaller than a pre-determined tolerance level (e.g.  $10^{-4}$ ); else, let  $P^{\{k+1\}} = P^{\{*\}}$ , set  $k = k + 1$  and go to step 4.
10. Obtain an approximately optimal exercising boundary based on  $P^{\{k\}}$ .
11. Generate a set of  $N_2$  ( $N_2 > N_1$ ) sample paths of electricity and fuel prices and evaluate the tolling contracts over these sample paths based on the optimal boundary obtained at step 10.

### 3.6 Hedging Tolling Contracts

If markets were complete, then the value function of a tolling agreement plus the cumulative payouts deposited into a bank should have a martingale representation with respect to the filtration generated by the Brownian motions and the compound Poisson process in (47) and (48), thus a perfect hedging strategy with continuous trading would exist. The electricity markets are inherently incomplete and the continuous trading is only an ideal assumption. Nevertheless, delta-hedging strategies derived with the continuous-trading assumption and implemented through discrete-trading still provide great practical value. In this section, I present a heuristic delta-hedging strategy for hedging a tolling contract.

As explained in section 3.3, tolling options in a tolling agreement closely resemble a strip of spark spread call options with maturity time spanning through the contract period. This observation prompts the idea of using delta positions derived for the European-style spark spread options to construct a hedging portfolio of a tolling contract.

To illustrate how the proposed hedging strategy works, I present an example of hedging a one-year tolling contract. First, I replicate the one-year tolling agreement by a portfolio of spark spread call options with maturity times ranging from 1-month to 12-month and strike heat rate equaling the plant's average heat rate. In the spark spread option portfolio, options of different maturities all have the same number of shares, which equal the capacity of the underlying plant. Also, I adjust the option positions every day to match the value of the option portfolio to that of the tolling agreement.

Under a typical delta-hedging scheme, to hedge an option on a generic underlying asset  $S_t$  with maturing time  $T$ , an instrument with payoff  $S_T$  at time  $T$  is needed. The instrument is usually the underlying asset itself if it is traded (e.g., in the case of hedging stock options). However, for an option on electricity, I cannot hold electricity to get  $S_T$  at time  $T$  since electricity is non-storable. Instead, an electricity futures contract with a payoff of  $S_T - F_E$  at time  $T$  is needed for hedging electricity options, where  $F_E$  is the electricity futures price. I use a futures contract and a bond to construct a synthetic security that is traded and pays out the electricity price  $S_T$  at time  $T$  for implementing delta-hedging for a tolling agreement. Specifically, each share of the synthetic security consists of one share of a futures contract and  $F_E$  shares of a riskless bond paying \$1 at time  $T$ . As the time- $T$  payoffs of the futures contract and the bond are  $S_T - F_E$  and  $F_E$ , respectively, they yield the synthetic security a combined payoff of  $S_T$ . The price of the synthetic security at time 0 is  $F_E \cdot e^{-rT}$  since the cost of entering into futures contracts is zero, where  $r$  is the risk-free rate.

I next construct a hedging portfolio out of a bank account and 24 synthetic securities consisting of 12 monthly electricity futures and 12 monthly nature gas futures over a time period of one year. For  $t \in [0, 1]$ , let  $V_t$  denote the remaining value of a tolling contract and  $M_t$  denote the month in which  $t$  falls. The hedging portfolio of  $V_t$  is denoted by  $(\vec{\delta}_E(t), \vec{\delta}_G(t), B_t)$  where  $\vec{\delta}_E(t) \equiv (\delta_E^{M_t}(t), \delta_E^{M_t+1}(t), \dots, \delta_E^{12}(t))$  represent the shares of synthetic securities based on electricity futures with maturity date  $T_{M_t}, T_{M_t+1}, \dots, T_{12}$  (while time-to-maturity ranges from 1-month to  $(13 - M_t)$ -month),  $(\delta_G^{M_t}(t), \dots, \delta_G^{12}(t))$  represent the shares of the corresponding synthetic natural gas contracts, and  $B_t$  represents the balance of a bank account. Let  $C_{t,T}$  denote the time- $t$  value of a spark spread call with a

payoff of  $\max(S_E(T) - H \cdot S_G(T), 0)$  at maturity  $T$  with  $H$  being the average heat rate of the power plant and  $T$  ranging from  $T_{M_t}$  to  $T_{12}$ . The shares of  $C_{t,T}$  ( $T = T_{M_t}, T_{M_t+1}, \dots, T_{12}$ ) to hold at time  $t$  for hedging  $V_t$  are obtained by solving for  $Q_t$  in (59).

$$V_t = Q_t \cdot \sum_{i=M_t}^{12} C_{t,T_i} \quad (59)$$

Therefore  $\delta_E^i(t)$  is given by

$$\delta_E^i(t) = Q_t \cdot \frac{\partial C_{t,T_i}}{\partial F_E^i} \quad \forall i = M_t, M_t + 1, \dots, 12 \quad (60)$$

where  $F_E^i$  is the electricity futures price of maturity  $T_i$  and  $C_{t,T_i}$  is calculated by the transform method in Deng [23]. The delta positions of the gas synthetic securities are obtained by solving (61) utilizing (60). That is, the total value of synthetic securities equal to the value of the tolling agreement less operating profit (loss) for each day.

$$Q_t \cdot C_{t,T_i} = e^{-r(T_i-t)} (\delta_E^i(t) F_E^i(t) + \delta_G^i(t) F_G^i(t)) \quad \forall i = M_t, M_t + 1, \dots, 12. \quad (61)$$

I re-balance the futures positions at the current futures prices at the end of each trading period  $\Delta t$ . Trading gains/losses get deposited or withdrawn from the bank account  $B_t$  (assuming no limit on the size of deposit and withdrawal). The balance of the bank account at the end of each trading period  $t$  is given by:

$$\begin{aligned} B_t = & \sum_{i=M_{t-1}}^{12} \{ \delta_E^i(t - \Delta t) (F_E^i(t) - F_E^{i-1}(t - \Delta t)) + \delta_G^i(t - \Delta t) (F_G^i(t) - F_G^{i-1}(t - \Delta t)) \} \\ & + \sum_{i=M_{t-1}}^{12} \{ \delta_E^i(t - \Delta t) F_E^i(t - \Delta t) e^{-r(T-t)} + \delta_G^i(t - \Delta t) F_G^i(t - \Delta t) e^{-r(T-t)} \} \\ & - \sum_{i=M_t}^{12} \{ \delta_E^i(t) F_E^i(t) e^{-r(T-t)} + \delta_G^i(t) F_G^i(t) e^{-r(T-t)} \} + B_{t-\Delta t} \cdot e^{r\Delta t}. \end{aligned} \quad (62)$$

By the end of the contract period, the balance  $B_{T_{12}}$  represents the cumulative hedging error.

### 3.7 Numerical Examples and Managerial Insights

To validate the valuation model, I apply it to a one-year tolling agreement on a hypothetical natural gas fired power plant with a capacity of 150 MW subject to the operational constraints and the maximum restart constraint discussed in section 3.3. Suppose the underlying plant sells wholesale power to the Electric Reliability Council of Texas (ERCOT)

region of the United States and purchases its fuel supply from the Henry Hub natural gas market. Thus I use ERCOT power price data and Henry Hub gas price data to estimate parameters for the stochastic price processes in (47) and (48).

I assume that the ramp-up cost rate function  $c(y)$  has the following form Deng and Oren [28].

$$c(y) = \underline{Q} * \underline{Hr} * e^y + M \quad (63)$$

where  $M$  is a constant. Table 4 summarizes the set of parameters characterizing the underlying power plant. In addition, I test our model for the maximal allowable number of restarts being three and six per year, respectively.

**Table 4:** Parameters for a Hypothetical Natural Gas Fired Power Plant

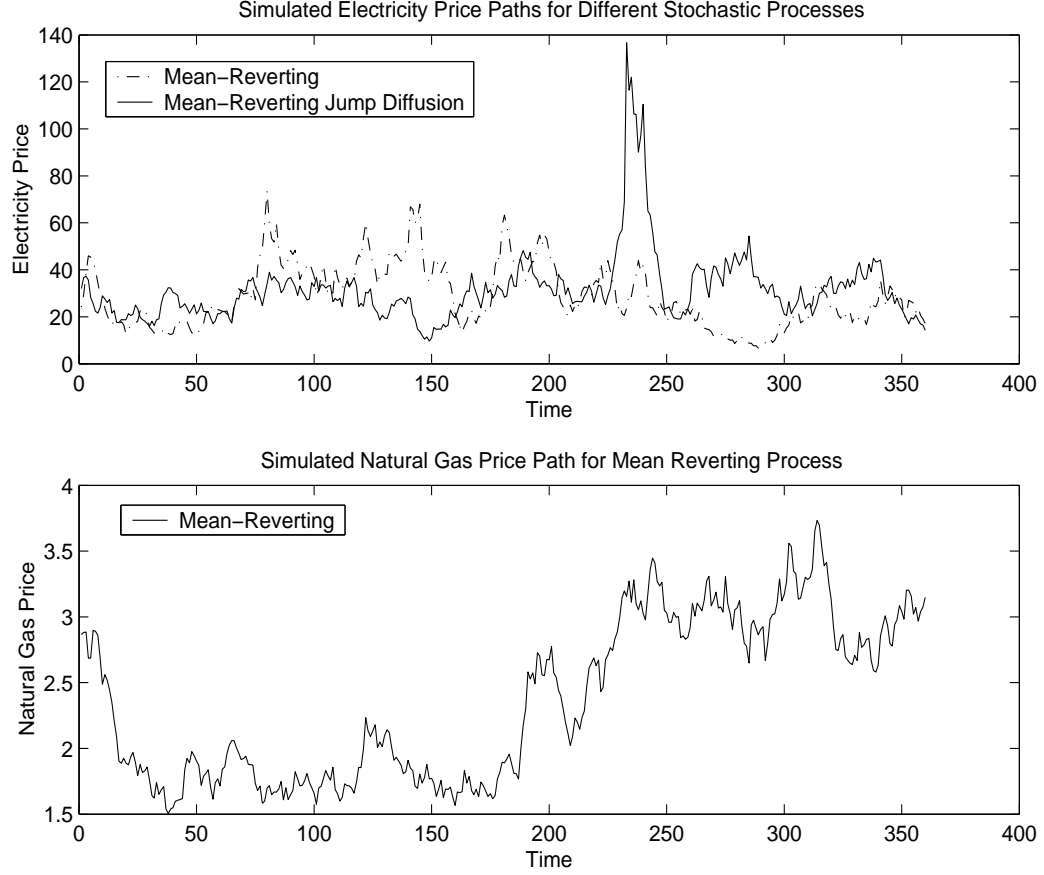
$C_{up}$	$C_{down}$	$\underline{Q} : \overline{Q}$	$\underline{Hr} : \overline{Hr}$	$M$	$r$	$D$
\$2000	\$1000	0.2 : 1	1.38 : 1	1	5%	1

I examine the effects of different electricity price modeling assumptions on the valuation of a tolling agreement using the two cases specified in section 3.3.2.

- Case 1: both power and gas price processes are mean-reverting (MR) processes.
- Case 2: the power price process is a mean-reverting jump-diffusion (MRJD) process and the gas price process is a MR process.

While both power and gas prices have the mean-reverting feature, power prices are much more volatile than other commodity prices (including gas price) as demonstrated by the salient spikes and jumps in the historical power prices. A MRJD process is a more realistic assumption for modeling power prices than a simple MR process. Case 1 serves as a benchmark case in which the price jumps are not explicitly modeled and case 2 offers the contrasting results with explicit modeling of price jumps. Tables 5 and 6 report parameters for the power and the gas price processes, respectively, which are estimated from the historical ECORT daily electricity prices and Henry Hub daily gas prices. The correlation coefficient  $\rho$  between the power price and the gas price is set to be the sample correlation

0.177 in both cases. The initial prices of electricity and natural gas are sampled from the historical data as \$34.7 per MWh and \$3 per MMBtu, respectively. Figure 8 plots three simulated sample paths for power and gas price models (47) and (48) with the estimated parameters in tables 5 and 6.



**Figure 8:** Simulated Sample Paths of Different Electricity and Gas Price Models

**Table 5:** Parameters for the Power Price Process

	$\mu$	$\sigma$	$\alpha$	$\phi$	$\bar{\kappa}$	$\gamma$
<i>MR</i>	3.5527	0.1507	0.0651	—	—	—
<i>MRJD</i>	3.5304	0.1299	0.0584	0.0281	0.0483	0.2566

I compute the value of the one-year tolling agreement for different levels of heat rate  $\overline{Hr}$ : 7.5, 10.5 and 13.5 MMBtu/MWh. Table 7 reports the estimates of the time-0 value of the tolling agreement and the corresponding standard errors based on 2000 Monte Carlo

**Table 6:** Parameters for the Gas Price Process

	$\mu$	$\sigma$	$\alpha$
<i>MR</i>	1.3638	0.0468	0.0087

sample paths of the power and gas price processes. The contract values in case 1 are slightly lower than those in case 2. A MRJD electricity price model is more realistic than a simple MR model, given the empirical features of power prices such as mean-reversion, jumps, and spikes (See Deng [23]). Therefore, I argue that the contract value obtained with the MRJD power price model is a better approximation to the fair value than the value obtained with the MR model.

To put our valuation model into perspective, several large energy merchant firms in the U.S. electricity industry paid approximately \$15 million per year in the early 2000s for tolling contracts in the ERCOT region written on a 150 MW power plant with an average heat rate of around 8.0 MMBtu/MWh, a similar cost structure to our hypothetical power plant, and with approximately 20 restarts allowed. Such a tolling premium is quite close to the estimated values in table 7. However, as I incorporate only a few of the major operational characteristics and constraints into our valuation model, the true value of a tolling agreement should be lower than our model-predicted value. Another important point is that the values in table 7 are computed under the assumption of underlying prices processes being stationary which means that the electricity price (or, fuel price) regime remains the same over time. However, if either of the price regimes were to change from the present state to a fundamentally different one in the future due to dramatic increases in energy demand (or, a flood of newly built generation capacity coming online), then the value of a tolling agreement would be significantly different from the value obtained under the stationary price assumption (this is further discussed in the conclusion section).

Figures 9 and 10 plot the optimal action regions for executing this tolling contract at the end of the third month and the sixth month over the 12-month contract period. Specifically, it is optimal for the contract holder to start taking electricity (i.e., exercising a tolling option,



**Table 7:** Estimated Values (in \$ millions) of the Tolling Agreement for Different Price Models and Heat Rates (Max\_Restarts N: 3 and 6; STD: standard error).

$N = 3 : \overline{Hr}$	7.5	8.0	10.5	13.5	$N = 6 : \overline{Hr}$	7.5	8.0	10.5	13.5
MR	15.02	14.94	8.09	4.06	MR	16.29	15.08	8.91	4.87
(STD-1)	0.28	0.33	0.27	0.18	(STD-1)	0.32	0.32	0.29	0.20
MRJD	15.40	15.18	8.33	4.11	MRJD	16.79	15.31	9.48	4.79
(STD-2)	0.32	0.34	0.28	0.17	(STD-2)	0.34	0.34	0.31	0.21

or turning on the plant) in the region to the south-east of the boundary line formed by the  $\times$ 's and stop taking electricity (i.e., terminating an exercised tolling option, or shutting down the plant) in the region to the north-west of the boundary line formed by the circles. I term these two regions as the “tolling” and the “no-tolling” regions. I choose 4 contract states (shown in the figures) to plot the corresponding “tolling” and “no-tolling” regions for the heat rates being 7.5 and 13.5 MMBtu/MWh. Turn-on and turn-off boundaries are clearly shaped in each plot.

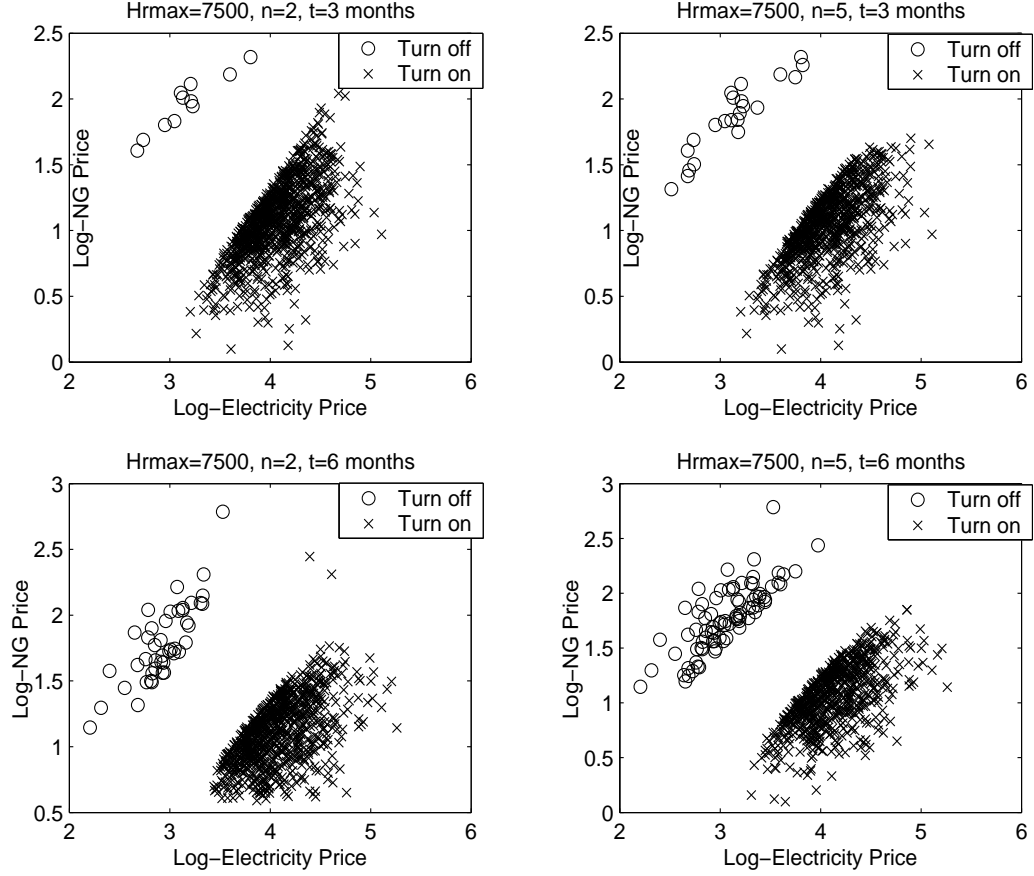
These computational results shed light on the structure of the optimal execution strategies for a tolling agreement and offer valuable insights for managing and operating such contracts. First of all, the optimal actions for operating and managing a tolling contract are separated into “tolling” and “no-tolling” regions in the plane of all possible price pairs of electricity and the fuel by some curves (or, boundaries). With this insight, a tolling contract holder knows that the optimal execution strategy of the contract is governed by some threshold curves in the form of certain functional relationships between the electricity price and the fuel price. Thus she or he can efficiently identify the optimal action regions for operational guidance by testing various relationships between the power price and the fuel price. I also observe that, as the remaining time of the contract gets shorter, both the “no-tolling” and the “tolling” regions get larger meaning that it is optimal for the contract holder to exercise the “tolling” and “shutting down” options more frequently as the time approaches contract expiration. On the other hand, the contract holder should be patient in determining whether to start taking electricity or to terminate an exercised tolling option

immediately in the early stage of the tolling contract, since there are only a limited number of re-start opportunities available. The two regions also get larger as the remaining number of re-starts gets larger since a larger number of re-starts reduces the opportunity cost of exercising a tolling option/shutdown option and leaves more flexibility with the contract holder in determining the best time to exercise a tolling or shutdown option.

All price pairs falling in between the “tolling” region and the “no-tolling” region constitute a “no action” band in the sense that, if a pair of the electricity price and the fuel price belongs to this band, the optimal action for the contract holder is to maintain status quo under current market conditions: either to continue taking the electricity output if under the obligation of an exercised tolling option, or to keep putting off the exercising of a tolling option if not under any tolling option’s obligation. Based on numerical experiments, I observe that the area of the no-action band shrinks as the remaining contract time gets shorter, and it expands as the number of remaining re-starts gets smaller. This implies that, as the tolling contract approaches expiration, the holder should be more active in determining which best action to take rather than passively sticking to its current operating state. On the other hand, when the number of remaining restarts gets smaller, the holder should be patient in determining the optimal action for now so as to leave the limited optionality for the best time to capture the most economic benefits.

The mean, standard deviation, and 90% confident interval of the cumulative hedging errors for case 2 are reported in table 8. I observe that while the mean and the standard deviation of hedging errors decrease as the heat rate increases, the percentage of the hedging error with respect to the tolling contract value decreases as the heat rate decreases. Namely, the percentage hedging error of the delta-hedging strategy for a tolling contract written on an efficient power plant (i.e., with low heat rate) is smaller than that for a contract written on an inefficient plant. This is as expected because the hedging strategy for a tolling contract is designed based on the hedging portfolio of a series of spark spread call options. For a tolling agreement on an efficient power plant, it can be well approximated by a series of spark spread call options.

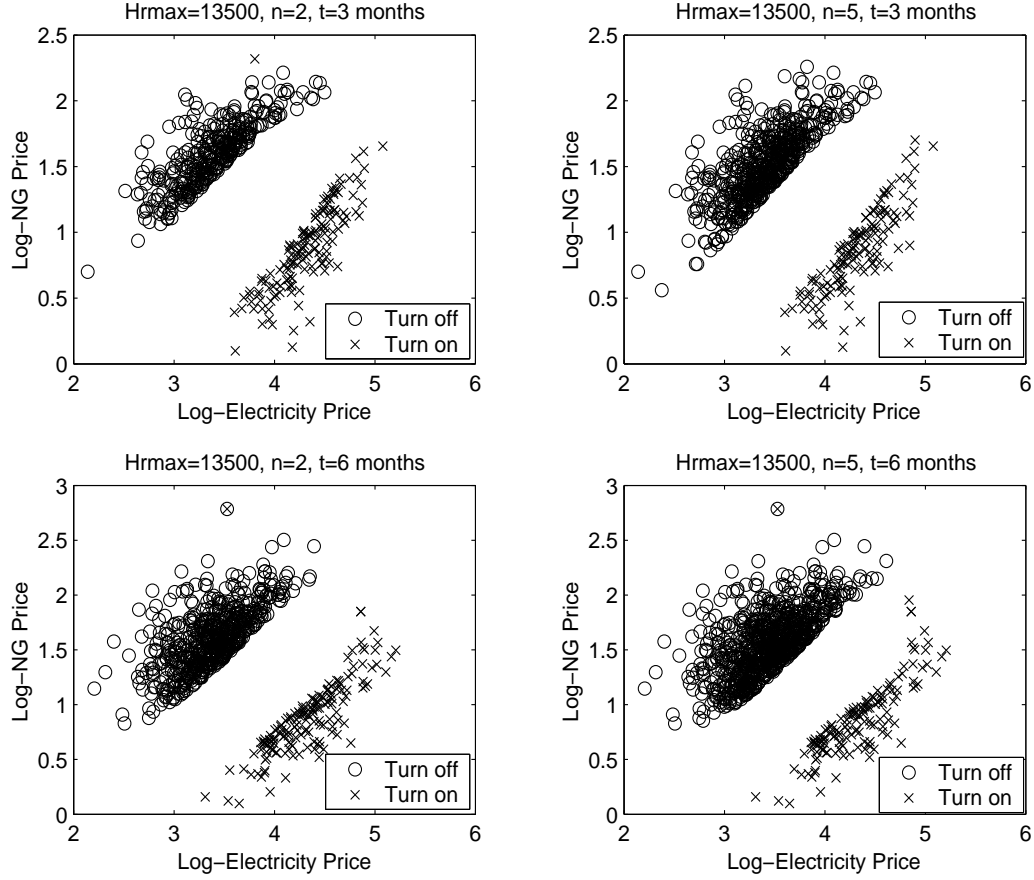
The two panels in figure 11 show the histograms of cumulative hedging errors, with



**Figure 9:** Optimal boundaries for heat rate 7.5

x-axis indicating the dollar amount for different heat rates with the maximal number of restarts being six. For heat rate being 7.5 and 13.5 MMBtu/MWh, the distributions of hedging errors skew to the right.

In table 9, I report the average values and standard deviations of a tolling agreement with various allowed number of restarts and  $\overline{Hr} = 7.5$  for case 2 for the two valuation methods described in section 3.4 and 3.5. For the purpose of comparison, I compute all values based on the same simulation set with 5,000 sample paths. In the implementation of the adaptive policy estimation method, I simulate another sample set with 100 sample paths to estimate parameters. Figures 12 and 13 plot the estimated policies for on-peak and off-peak at different times with three restarts remaining and six restarts remaining respectively. One could notice that the turn-on and turn-off boundaries for  $t = 0.5$  year intersect with each other at some place with a very low electricity price. It is primarily due



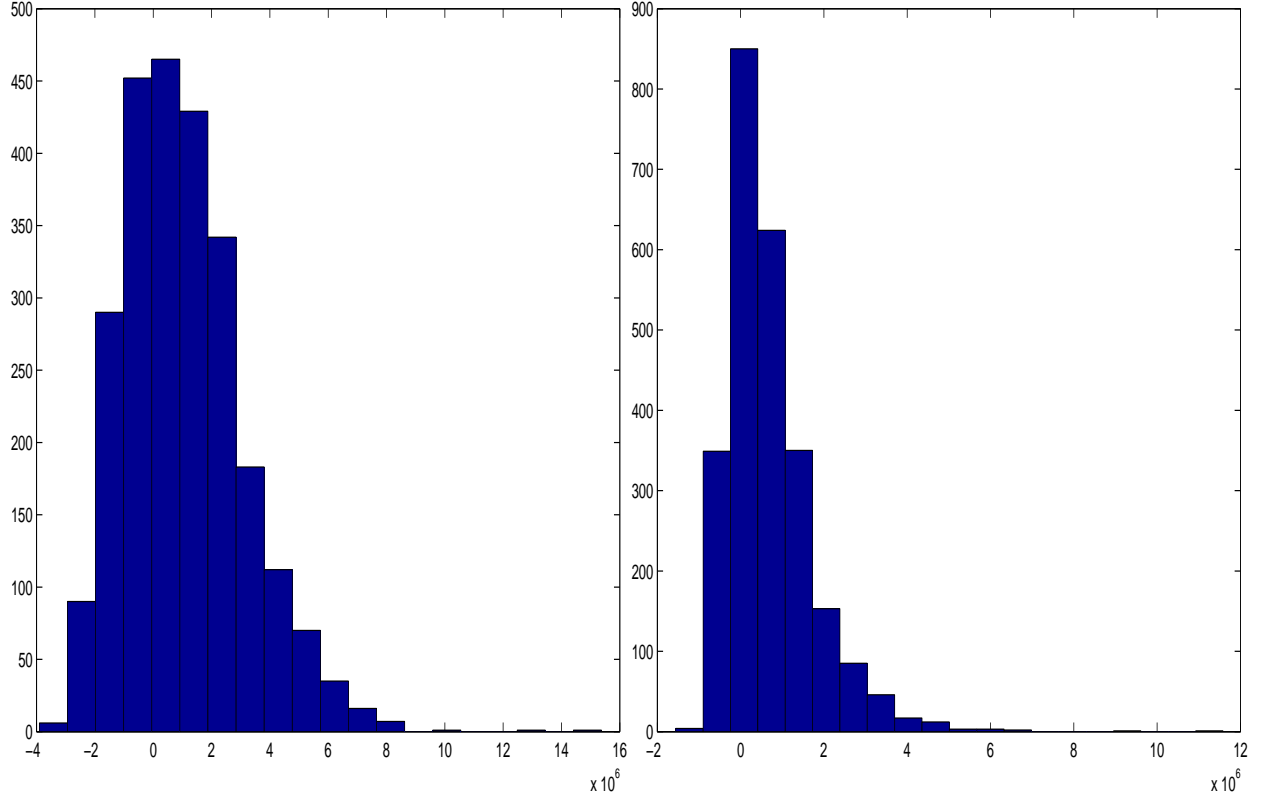
**Figure 10:** Optimal boundaries for heat rate 13.5

to the fact that I do not observe such low prices in our simulation study, and therefore can not estimate the policies for low electricity prices accurately. If such low electricity prices are observed, I feel that a piecewise linear policy function will be better than the linear policy function. Figures 14 and 15 plot the slopes and intercepts of the turn-on/turn-off boundaries as functions of time. The slope of the turn-off boundary first increases as time elapses and then flattens out. On the other hand, the slope of the turn-on boundary are always flat. The intercepts of the turn-off and turn-on boundaries both decrease initially in time and then become flat after a period of time elapses. With an increasing slope and decreasing intercept of the turn-off boundary, the turn-off region expands at the low fuel price levels and shrinks at the high fuel price levels in the early stage of the tolling contract. Since the spark spread on the turn-off boundary is negative, the tolling contract holder can tolerate more negative spark spreads in the early stage. As the time under contract elapses,

**Table 8:** Statistics of Hedging Errors for Case 2 (Max\_Restarts N: 3 and 6)

$N = 3 : \overline{Hr}$	7.5	10.5	13.5
<i>Mean</i>	1.03	1.02	0.66
<i>Percentage</i> (%)	6.69	12.27	16.06
<i>STD</i>	2.01	1.56	1.02
$Conf_{5\%-95\%}$	(-1.79, 4.67)	(-0.83, 3.98)	(-0.42, 2.60)
$N = 6 : \overline{Hr}$	7.5	10.5	13.5
<i>Mean</i>	1.03	0.95	0.69
<i>Percentage</i> (%)	6.13	10.02	14.41
<i>STD</i>	2.06	1.48	1.05
$Conf_{5\%-95\%}$	(-1.86, 4.90)	(-0.88, 3.72)	(-0.46, 2.65)

the contract holder would tolerate less negative spreads and thus would turn off the plant more frequently when the start-up and ramp up costs, which depends on the fuel prices, are low. This results in the expanded turn-off region at the low fuel price levels. On the other hand, the marginal value of one restart decreases at the high fuel price levels (high ramp up costs) as time approaches the contract termination date, thus, leading to a shrunk turn-off region. After a period of contract time elapses, the slope of the turn-off boundary becomes flat first and thus the turn-off region become larger with the decreasing intercept. Finally, when the intercept becomes flat, the turn-off region will not change until the maturity. This is because the contract holder does not want to leave the restart opportunities wasted as time approaches the contract termination and thus will turn off the plant more frequently. When no restart is left, the turn-off region, of course, will not change any more. Similarly, the turn-on region is reduced in the early stage of the contract is smaller than that in the late stage of the contract because the contract holder would like to have more positive spark spreads as time approaches the contract termination considering the start up cost and ramp up cost.



**Figure 11:** Histogram of Hedge Errors:  $Hr = 7.5$  (left) vs.  $Hr = 13.5$  (Right)

### 3.8 Conclusion and Future Work

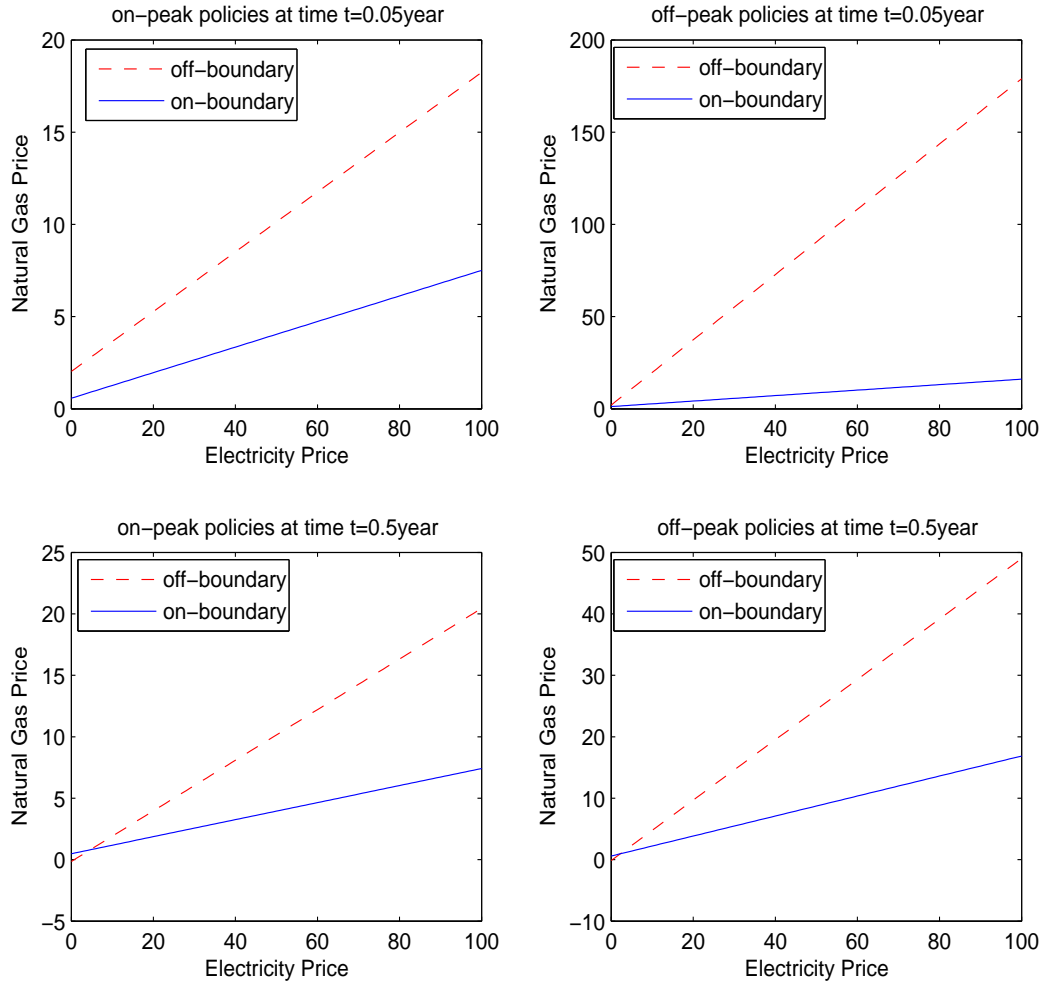
Electricity tolling agreements, as well as other structured transactions, have played important roles in facilitating risk-sharing and risk-mitigation among independent power producers, utility companies, and unregulated power marketers in the restructured power industry. As power markets continue to evolve, market information and variables needed for pricing these complex structured transactions will become more and more transparent. However, at the present time, I need to work with assumptions, that are plausible but yet to be empirically justified in investigating the problem of pricing and hedging customized electricity contracts such as tolling agreements. Under these assumptions, I formulate a real option based valuation model and solve it by two methods—the least-squares regression method and the adaptive policy estimation method. Both methods yield a fairly accurate approximation to the market value of a tolling agreement as gauged by the limited available market transaction data.

**Table 9:** Comparison of valuation by least-square Monte Carlo and adaptive policy search. (All values are in \$ millions)

<i>#of Restarts</i>	1	3	5	8	10
<i>LSMethod</i>	14.514	14.916	15.054	15.176	15.233
( <i>STD</i> )	(0.3516)	(0.3410)	(0.3370)	(0.3334)	(0.3317)
<i>PolicyMethod</i>	14.394	14.964	15.202	15.314	15.329
( <i>STD</i> )	(0.0556)	(0.0718)	(0.0794)	(0.0765)	(0.0778)

Through numerical examples, I also examine the effects of different power price assumptions on the tolling contract valuation. The simple mean-reverting power price model results in a slightly lower tolling premium than does the mean-reverting jump-diffusion power price model. The well-known empirical features of electricity prices make the jump-diffusion assumption a more realistic choice for pricing tolling agreements. One other crucial factor affecting the valuation is the stationarity assumption on the underlying price processes. Electricity prices are indeed non-stationary due to seasonality effects, fundamental changes in electricity supply and demand, evolutions in market designs, and other modifications to regulatory policies. Tolling contracts signed in the late 1990s turned out to be significantly overvalued as they were priced under the assumption that electricity prices would fluctuate in a high-price regime which was made based on the limited market price data at that time. The risk of a regime shift in electricity prices from a high-price regime to a low-price one was not properly incorporated, thus the overvaluation. To mitigate such over- or under-valuation risks, one needs to account for non-stationarity in modeling the underlying prices, particularly when valuing a contract with a long horizon such as 10 to 15 years.

One point worth mentioning is that the real options valuation formulation is applicable under quite general price modeling assumptions. If the underlying prices are non-stationary with jumps, most numerical schemes such as the lattice approach and the partial differential equation approach would encounter difficulties in solving real options valuation problems. In such cases, the value function approximation scheme and the adaptive policy estimation method based on Monte Carlo simulation are two viable approaches for solving the valuation

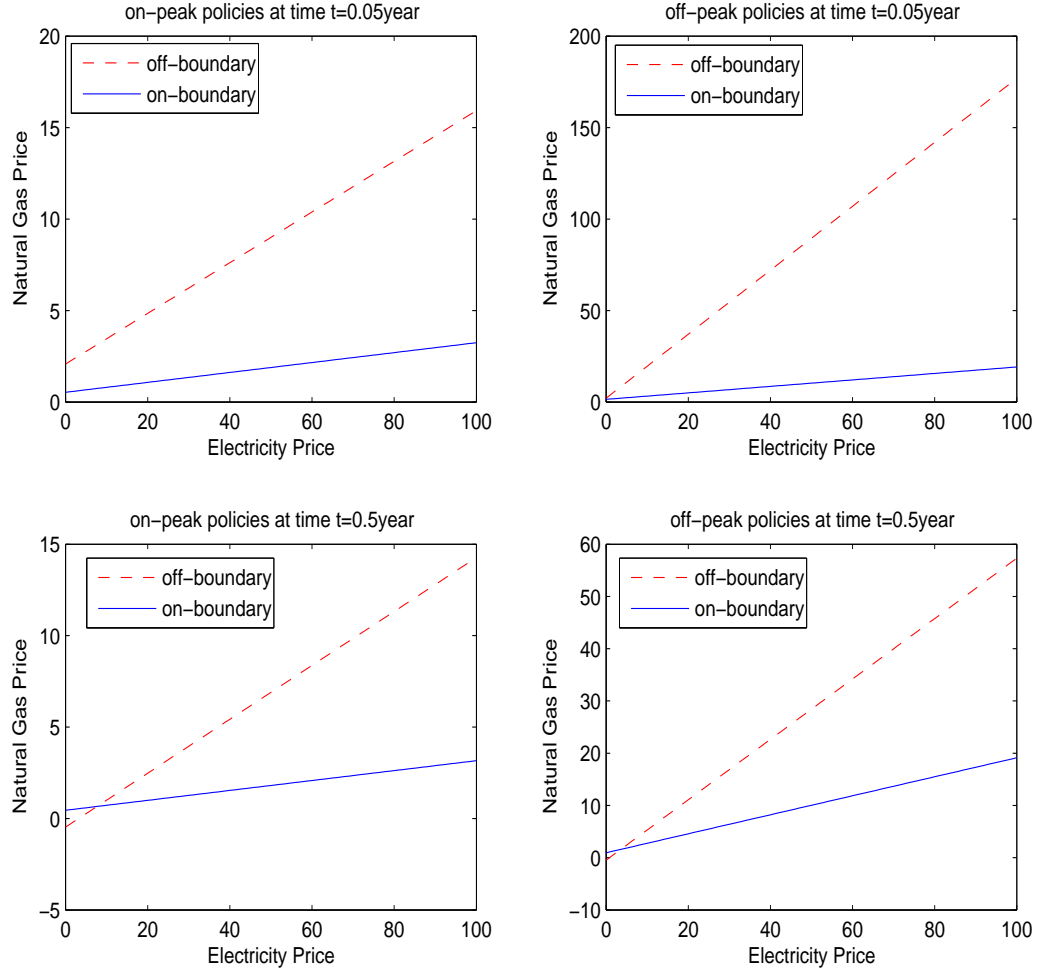


**Figure 12:** Estimated policies for 3 restarts remaining

problem.

In regard to the other important issue on hedging a tolling agreement, I propose a heuristic delta-hedging scheme by utilizing futures contracts. Starting with the premium of a tolling contract, I construct a portfolio for hedging (or replicating) the contract using electricity and generating fuel futures and a riskless bank account. The hedging position in electricity futures is calculated through the closed-form pricing formula of a spark spread call option, and the position in generating fuel futures is obtained by matching the value of the hedging portfolio to that of the tolling contract. I then continuously re-balance the hedging portfolio by adjusting the number of shares of each security held. In the numerical

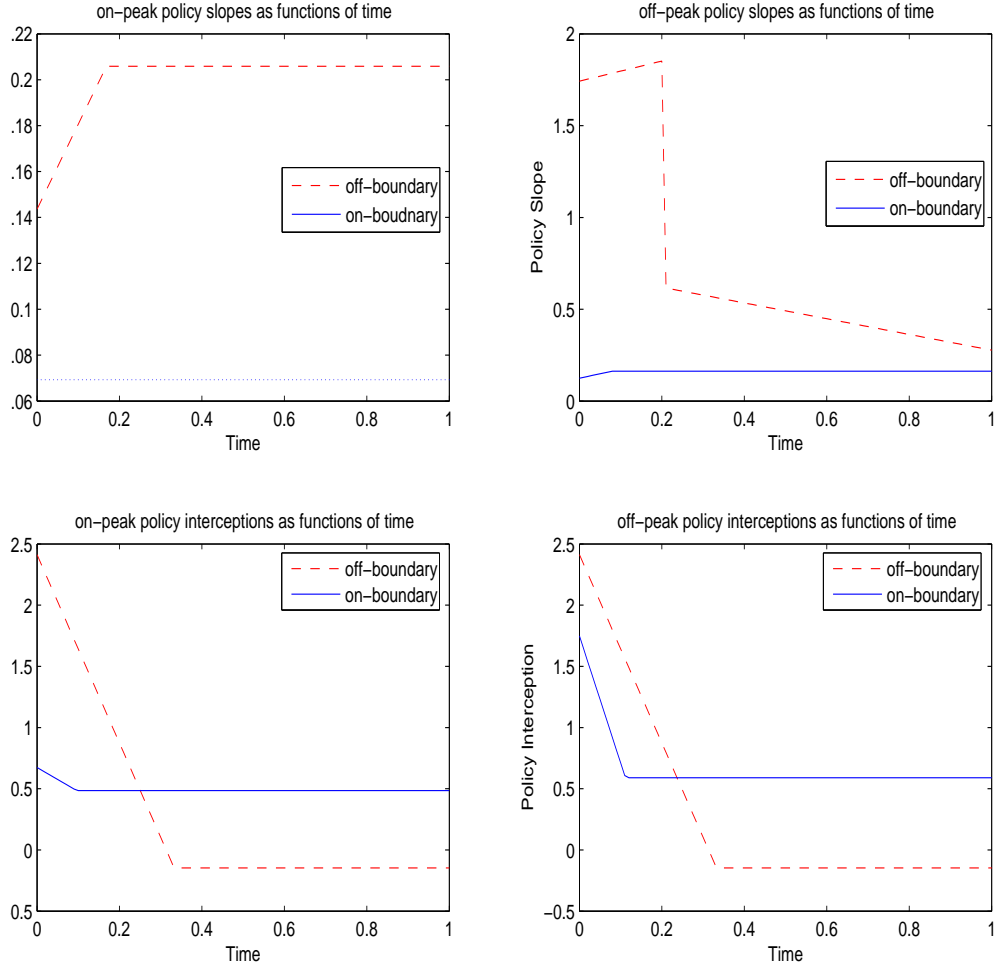




**Figure 13:** Estimated policies for 6 restarts remaining

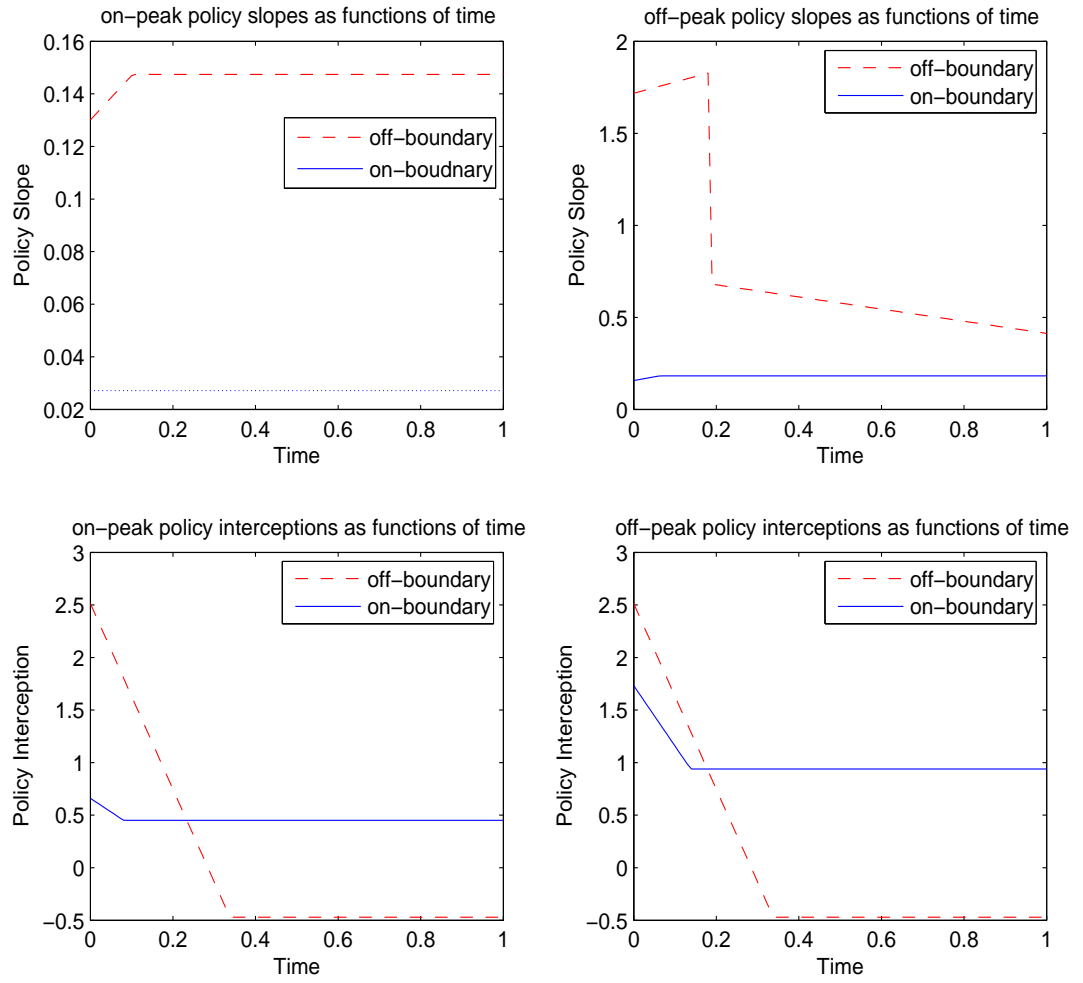
examples, this straightforward hedging scheme results in an approximate 6% hedging error with respect to the tolling contract value for an efficient power plant (i.e. heat rate is 7.5), and a 14 ~ 16% hedging error for an inefficient power plant (i.e. heat rate is 13.5). The less/more efficient a power plant is or the less/more frequent the restarts are allowed, the larger/smaller percentage error the hedging scheme yields.

As for future work, I plan to carry out rigorous empirical tests on one major assumption of the tolling contract valuation model: the specification of a power price model. Another fruitful direction deserving further pursuit is the search for a more efficient and accurate



**Figure 14:** Policy slopes and interceptions as functions of time for 3 restarts remaining

hedging scheme for a tolling contract, especially when the underlying power plant is inefficient meaning that it has a very large operating heat rate. All of the discussion in this chapter is under the real measure. The purpose is to present the simulation-based real option valuation framework for pricing a tolling contract. The underlying processes could be under either real measure or risk neutral measure. I use the processes under real measure to illustrate the valuation framework. In the future, I could make use of the existing versions of both mean-reverting and mean-reverting jump diffusion under risk neutral measure to value a tolling contract. For example, a mean-reverting process under real measure has a parameter vector denoted by  $\theta$  and the one under risk neutral measure has a parameter



**Figure 15:** Policy slopes and interceptions as functions of time for 6 restarts remaining vector denoted by  $\theta^*$ . What I need to do is only to estimate  $\theta^*$  based on market risk premiums.

## CHAPTER IV

# THE CONDITIONAL VALUE-AT-RISK UNDER A GARCH PRICE MODEL WITH HEAVY TAILED INNOVATIONS

### *4.1 Introduction*

Heavy tails and heteroscedasticity are two important features of empirical financial return distribution. ARCH and GARCH models were proposed to model the conditional second moments and long term dependence structure of financial return series. However, the classical ARCH/GARCH model has a normal innovation and therefore captures the heavy tails of return series. In this chapter, I will deal with a GARCH(1,1) model with heavy tailed innovation.

As a key risk measure in practice, the value-at-risk (VaR) is important for most financial institutions. The VaR is mathematically defined as a quantile of the return distribution of a portfolio. A more interesting risk measure is the conditional VaR, which measures the worst possible financial loss to a portfolio over a fixed time horizon for a given confidence (or probability) level conditional on current market information. While a large number of efforts have been focused on producing new and better conditional VaR estimates, two sources of errors may affect the estimation accuracy significantly: model mis-specification error and estimation error due to the inherent noise in the data. I address these two problems by considering non-parametric heavy-tailed distributions for the conditional innovations of a GARCH model. I then obtain the confidence intervals for the conditional VaR estimators of the heavy-tailed GARCH model. The knowledge of the confidence interval of the conditional VaR can be highly valuable in applications such as setting prudent capital reserve requirements for banks and conservative trading limits for traders or evaluating corporate self-insurance exposures by providing upper and lower bounds, rather than a single point

estimate, of the VaR estimator at a certain confidence level.

I have two main objectives in the chapter. First of all, I derive the limiting distribution of the extreme conditional VaR estimator in McNeil and Frey [68]. Instead of working within the framework of generalized Pareto distribution as in McNeil and Frey [68], I deal with the heavy-tailed innovations. In particular, besides the heavy-tailed feature, no specific parametric distributional assumptions on the GARCH innovations are imposed. A major advantage of this non-parametric approach is that it is applicable regardless of the true data-generating mechanism of the GARCH innovations, as long as it has heavy tails. As pointed out by Rachev and Mittnik [71], one weakness of the VaR methodology comes from model (mis-)specification risk. With a non-parametric model, I can mitigate this model risk. Another advantage is that in addition to a VaR estimator, I can provide a VaR interval so that financial institutions can calculate the risks of loss exceeding the upper boundary of the VaR interval. This information amends the ability of quantifying and controlling risks. Some existing work proposes to use the bootstrap method for constructing confidence intervals of a conditional VaR estimator (for instance, Dowd [31] and Christoffersen and Goncalves [17]). Note that the bootstrap method is computationally intensive because it requires repetitively solving non-linear optimizations in fitting GARCH models. I then estimate confidence intervals of a conditional VaR estimator by the traditional normal approximation method and the data tilting method.

The rest of the chapter is organized as follows: in section 4.2, I study the asymptotic behavior of the conditional VaR estimator by deriving its limiting distribution and give the normal approximation confidence interval estimation. I perform a simulation study, test our approach on electricity spot prices data in section 4.3 and implement the bootstrap method in Christoffersen and Goncalves [17] for the purpose of comparison, while proofs are given in appendix.

## 4.2 Model Specification and Estimation Methodology

Suppose the data generating process for observations  $\{X_t : t = \dots, -1, 0, 1, 2, \dots, n, \dots\}$  follows a GARCH( $p, q$ ) model, namely,

$$X_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = c + \sum_{i=1}^p b_i X_{t-i}^2 + \sum_{j=1}^q a_j \sigma_{t-j}^2 \quad (64)$$

where  $c > 0$ ,  $b_1 \geq 0, \dots, b_p \geq 0$ ,  $a_1 \geq 0, \dots, a_q \geq 0$  are constants,  $\{\epsilon_t\}$  are a sequence of independent identically distributed random variables with mean 0 and variance 1 (i.e.,  $IID(0, 1)$ 's), and  $\epsilon_t$  is independent of  $\{X_{t-k}, k \geq 1\}$  for all  $t$ . Further assume that (64) uniquely defines a strictly stationary process with  $EX_t^2 < \infty$ , i.e.,

$$\sum_{i=1}^p b_i + \sum_{j=1}^q a_j < 1. \quad (65)$$

The  $100\alpha$  ( $0 < \alpha < 1$ ) percent one-step ahead conditional Value-at-Risk, based on observations  $\{X_1, \dots, X_n\}$ , is defined as

$$x_{\alpha, n} = \inf\{x : P(X_{n+1} \leq x | X_{n+1-k}, k \geq 1) \geq \alpha\}.$$

It is a straightforward derivation from (64) that  $x_{\alpha, n} = \sigma_{n+1} x_{\alpha}^0$  where  $x_{\alpha}^0$  is the  $100\alpha\%$  quantile of  $\epsilon_{n+1}$ . Our aim is to construct a confidence interval for the extreme conditional quantile  $x_{\alpha, n}$  (i.e.,  $\alpha = \alpha(n)$  tends to zero or one as  $n \rightarrow \infty$ ) through deriving the limiting distribution of an estimator of  $x_{\alpha, n}$  and then applying two interval estimation methods.

### 4.2.1 Point Estimation

In this subsection, I study the asymptotic behavior of an estimator for  $x_{\alpha, n}$  used in McNeil and Frey [68] under the assumption that  $\epsilon_t$  in (64) has heavy tails. Specifically, the distribution function  $G$  of  $\epsilon_t$  satisfies

$$\lim_{x \rightarrow \infty} \frac{1 - G(xy)}{1 - G(x)} = y^{-\gamma} \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{G(-x)}{1 - G(x)} = d \quad (66)$$

for all  $y > 0$ , where  $\gamma > 2$  is made to ensure that  $E\epsilon_t^2 < \infty$  and  $d$  is some constant in  $[0, \infty)$ .

Note that (65) implies that

$$\begin{aligned} \sigma_t^2(a, b, c) &= \frac{c}{1 - \sum_{j=1}^q a_j} + \sum_{i=1}^p b_i X_{t-i}^2 \\ &\quad + \sum_{i=1}^p b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^q \cdots \sum_{j_k=1}^q a_{j_1} \cdots a_{j_k} X_{t-i-j_1-\dots-j_k}^2, \end{aligned}$$

where  $a = (a_1, \dots, a_q)$  and  $b = (b_1, \dots, b_p)$ . In practice I replace the above expression by a truncated version

$$\begin{aligned}\tilde{\sigma}_t^2(a, b, c) &= \frac{c}{1 - \sum_{j=1}^q a_j} + \sum_{i=1}^p b_i X_{t-i}^2 + \sum_{i=1}^p b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^q \dots \sum_{j_k=1}^q a_{j_1} \dots a_{j_k} \\ &\quad \times X_{t-i-j_1-\dots-j_k}^2 I(t-i-j_1-\dots-j_k \geq 1)\end{aligned}$$

where  $I(\cdot)$  is an indicator function.

Define

$$L_\nu(a, b, c) = \sum_{t=\nu}^n \{X_t^2 / \tilde{\sigma}_t^2(a, b, c) + \log \tilde{\sigma}_t^2(a, b, c)\},$$

where  $\nu = \nu(n) \rightarrow \infty$  and  $\nu/n \rightarrow 0$  as  $n \rightarrow \infty$ . Then the quasi maximum likelihood estimator of  $(a, b, c)$  is defined as

$$(\hat{a}, \hat{b}, \hat{c}) = \operatorname{argmin}_{(a,b,c)} L_\nu(a, b, c).$$

Set

$$\lambda_n = \begin{cases} \inf\{\lambda > 0 : nP(\epsilon_t^2 \geq \lambda) \leq 1\} & \text{if } 2 < \gamma < 4 \\ \inf\{\lambda > 0 : nE(\epsilon_t^4 I(\epsilon_t^2 \leq \lambda)) \leq \lambda^2\} & \text{if } \gamma \geq 4. \end{cases}$$

Then, it follows from Hall and Yao [47] that

$$\hat{a} - a = O_p(n^{-1}\lambda_n), \quad \hat{b} - b = O_p(n^{-1}\lambda_n), \quad \hat{c} - c = O_p(n^{-1}\lambda_n).$$

Thus,  $\epsilon_t$  can be estimated by  $\hat{\epsilon}_t = X_t / \tilde{\sigma}_t(\hat{a}, \hat{b}, \hat{c})$  for  $t = \nu, \dots, n$ . Next I use  $\hat{\epsilon}_t$ 's to estimate  $x_\alpha^0$  as follows. I only deal with the case  $\alpha = \alpha(n) \rightarrow 1$  as  $n \rightarrow \infty$ .

Let  $\hat{\epsilon}_{m,1} \leq \dots \leq \hat{\epsilon}_{m,m}$  denote the order statistics of  $\hat{\epsilon}_\nu, \dots, \hat{\epsilon}_n$  with  $m \equiv n - \nu + 1$ . Then  $\gamma$  can be estimated by the Hill estimator

$$\hat{\gamma} = \left\{ \frac{1}{k} \sum_{i=1}^k \log \frac{\hat{\epsilon}_{m,m-i+1}}{\hat{\epsilon}_{m,m-k}} \right\}^{-1},$$

where  $k = k(m) \rightarrow \infty$  and  $k/m \rightarrow 0$  as  $n \rightarrow \infty$  (see Hill (1975)). Replacing  $x$ ,  $1 - G(x)$ ,  $\gamma$  in (66) by  $\hat{\epsilon}_{m,m-k}$ ,  $\frac{1}{m} \sum_{i=1}^n I(\hat{\epsilon}_i > x)$  and  $\hat{\gamma}$ , respectively, I have  $1 - G(y\hat{\epsilon}_{m,m-k}) \sim \frac{k}{m} y^{-\hat{\gamma}}$ . Since  $1 - G(x_\alpha^0) = 1 - \alpha$ , I solve  $\frac{k}{m} y^{-\hat{\gamma}} = 1 - \alpha$  to obtain  $y = (1 - \alpha)^{-1/\hat{\gamma}} (\frac{k}{m})^{1/\hat{\gamma}}$ , i.e.,  $x_\alpha^0 \sim y\hat{\epsilon}_{m,m-k}$ . So I estimate  $x_\alpha^0$  by

$$\hat{x}_\alpha^0 = (1 - \alpha)^{-1/\hat{\gamma}} \left(\frac{k}{m}\right)^{1/\hat{\gamma}} \hat{\epsilon}_{m,m-k},$$

i.e.,

$$\hat{x}_{\alpha,n} = \tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{x}_{\alpha}^0$$

is an estimator of  $x_{\alpha,n}$ .

Let  $U(x)$  denote the inverse function of  $\frac{1}{1-G(x)}$ . Suppose there exists some function  $A(x) \rightarrow 0$ , as  $x \rightarrow \infty$ , such that

$$\lim_{x \rightarrow \infty} \frac{U(xy)/U(x) - y^{1/\gamma}}{A(x)} = y^{1/\gamma} \frac{y^{\rho} - 1}{\rho}, \quad (67)$$

for all  $y > 0$ , where  $\rho < 0$ .

The following result characterizes the limiting distribution of the estimator  $\hat{x}_{\alpha,n}$ .

**Theorem 4.2.1** *Suppose (64), (65), (66), (67) and the conditions in Theorem 2.2 of Hall and Yao [47] hold. Assume*

$$k = k(m) \rightarrow \infty, k/m \rightarrow 0, \sqrt{k}A(m/k) \rightarrow 0, n^{-1}\lambda_n/A(m/k) \rightarrow 0, \log(\frac{k}{m(1-\alpha)})/\sqrt{k} \rightarrow 0$$

as  $n \rightarrow \infty$ . Then

$$\frac{\hat{\gamma}\sqrt{k}}{|\log(k/(m(1-\alpha)))|} \left\{ \frac{\hat{x}_{\alpha,n}}{x_{\alpha,n}} - 1 \right\} \xrightarrow{d} N(0, 1),$$

i.e.,

$$\frac{\hat{\gamma}\sqrt{k} \log(\hat{x}_{\alpha,n}/x_{\alpha,n})}{|\log(k/(m(1-\alpha)))|} \xrightarrow{d} N(0, 1).$$

**Remark 1.** As in Peng and Yao [80], I can re-parameterize the model (64) in such a way that the median of  $\epsilon_t^2$  is equal to 1 while keeping  $E(\epsilon_t) = 0$  unchanged. Under this new parameterization the parameters  $c$  and  $b_1, \dots, b_p$  differ from those in the old setting by a common positive constant factor while the parameters  $a_1, \dots, a_q$  remain unchanged. More importantly, the estimator  $\hat{x}_{\alpha,n}$  remains the same, but now the parameters can be estimated with convergence rate  $n^{-1/2}$  whenever  $E\epsilon_t^4 = \infty$  or  $< \infty$ . Therefore, with this parameter estimation, the condition  $n^{-1}\lambda_n/A(m/k) \rightarrow 0$  in Theorem 1 can be removed.



### 4.2.2 Interval Estimation

In this subsection I propose two methods to construct confidence intervals for the conditional VaR  $x_{\alpha,n}$  as follows.

**Method I: Normal approximation method.** Based on Theorem 1 above, a confidence interval with level  $\beta$  for  $x_{\alpha,n}$  is

$$I_{\beta}^n = (\hat{x}_{\alpha,n} \exp\{-z_{\beta} |\log \frac{k}{m(1-\alpha)}| / (\hat{\gamma}\sqrt{k})\}, \quad \hat{x}_{\alpha,n} \exp\{z_{\beta} |\log \frac{k}{m(1-\alpha)}| / (\hat{\gamma}\sqrt{k})\}),$$

where  $z_{\beta}$  satisfies  $P(|N(0,1)| \leq z_{\beta}) = \beta$ .

**Method II: Data tilting method.** The general data tilting method was proposed by Hall and Yao [46] to tilt time series data, and it includes the empirical likelihood method as a special case in general. The empirical likelihood method, introduced in Owen [76], [77], is a nonparametric approach for constructing confidence regions. Like the bootstrap and jackknife methods, the empirical likelihood method does not need to specify a family of distributions for the data. One of the advantages of the empirical likelihood is that it enables the shape of a region, such as the degree of asymmetry in a confidence interval, to be determined automatically by the sample. In certain regular cases, empirical likelihood based confidence regions are Bartlett correctable; see Hall and La Scala [45] and DiCiccio et al. [29]. For a more complete disclosure of recent references and development I refer to a book by Owen [78]. As a generalization of the empirical likelihood method, the data tilting method not only has all of those nice properties of the empirical likelihood method, but also admits a wide range of distance functions. Recently Peng and Qi [79] applied the data tilting method in Hall and Yao [46] to construct a confidence interval for the high quantile of a heavy tailed distribution based on iid observation. Here I apply the data tilting method in Peng and Qi [79] to the estimated innovations as follows.

Define  $\delta_i = I(\hat{\epsilon}_i \geq \hat{\epsilon}_{m,m-k})$ . First, for any fixed  $w = (w_{\nu}, \dots, w_n)$  such that  $w_i \geq 0$  and  $\sum_{i=\nu}^n w_i = 1$ , I solve

$$(\hat{\gamma}(w), \hat{c}(w)) = \operatorname{argmin}_{(\gamma, c)} \sum_{i=\nu}^n w_i \log((c\gamma\hat{\epsilon}_i^{-\gamma-1})^{\delta_i} (1 - c\hat{\epsilon}_{m,m-k}^{-\gamma})^{1-\delta_i}).$$

This results in

$$\hat{\gamma}(w) = \frac{\sum_{i=\nu}^n w_i \delta_i}{\sum_{i=\nu}^n w_i \delta_i (\log \hat{\epsilon}_i - \log \hat{\epsilon}_{m,m-k})}$$

and

$$\hat{c}(w) = \hat{\epsilon}_{m,m-k}^{\hat{\gamma}(w)} \sum_{i=\nu}^n w_i \delta_i.$$

Define

$$D_l(w) = \begin{cases} (l(1-l))^{-1} (1 - m^{-1} \sum_{i=\nu}^n (mw_i)^l) & \text{if } l \neq 0, 1 \\ -m^{-1} \sum_{i=\nu}^n \log(mw_i) & \text{if } l = 0 \\ \sum_{i=\nu}^n w_i \log(mw_i) & \text{if } l = 1. \end{cases}$$

Next, solve

$$(2m)^{-1} L(x_{\alpha,n}) = \min_w D_l(w)$$

subject to

$$w_i \geq 0, \sum_{i=\nu}^n w_i = 1, \hat{\gamma}(w) \log(x_{\alpha,n}/(\tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{\epsilon}_{m,m-k})) = \log((\sum_{i=\nu}^n w_i \delta_i)/(1-\alpha)).$$

Here I only consider the case  $l = 1$  since other cases are similar, and the case  $l = 1$  gives good robustness properties. Put

$$A_1(\lambda_1) = 1 - \frac{m-k}{m} e^{-1-\lambda_1}, \quad A_2(\lambda_1) = A_1(\lambda_1) \frac{\log(x_{\alpha,n}/(\tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{\epsilon}_{m,m-k}))}{\log(A_1(\lambda_1)/(1-\alpha))}.$$

Then, by the standard method of Lagrange multipliers, I have

$$w_i = \begin{cases} \frac{1}{m} e^{-1-\lambda_1}, & \text{if } \delta_i = 0 \\ \frac{1}{m} \exp\{-1 - \lambda_1 + \lambda_2 (\frac{\log(x_{\alpha,n}/(\tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{\epsilon}_{m,m-k}))}{A_2(\lambda_1)} - \frac{1}{A_1(\lambda_1)} - \frac{A_1(\lambda_1)}{A_2^2(\lambda_1)} \log(\hat{\epsilon}_i/\hat{\epsilon}_{m,m-k}) \log(x_{\alpha,n}/(\tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{\epsilon}_{m,m-k})))\} & \text{if } \delta_i = 1, \end{cases}$$

where  $\lambda_1$  and  $\lambda_2$  satisfy

$$\sum_{i=\nu}^n w_i = 1 \quad \hat{\gamma}(w) \log(x_{\alpha,n}/(\tilde{\sigma}_{n+1}(\hat{a}, \hat{b}, \hat{c})\hat{\epsilon}_{m,m-k})) = \log(\sum_{i=\nu}^n w_i \delta_i/(1-\alpha)).$$

**Theorem 4.2.2** *Under the conditions of Theorem 1,*

$$L(x_{\alpha,n}^0) \xrightarrow{d} \chi^2(1)$$

as  $n \rightarrow \infty$ , where  $x_{\alpha,n}^0$  denotes the true value of  $x_{\alpha,n}$ .

Based on this theorem, a confidence interval with level  $\beta$  for  $x_{\alpha,n}^0$  can be constructed as

$$I_{\beta}^t = \{x_{\alpha,n} : L(x_{\alpha,n}) \leq u_{\beta}\},$$

where  $u_{\beta}$  is the  $\beta$ -level critical point of  $\chi^2(1)$ .

### ***4.3 Simulation Study and Application***

In this section, I investigate the finite sample behavior of our methods in constructing confidence intervals for the extreme conditional Value-at-Risk. And I also apply the methodology to a real data set taken from energy (e.g., electricity) markets.

#### **4.3.1 Simulation Study**

I draw 1,000 samples of size 2000 from GARCH(1,1) model with  $c = 1.0, b_1 = 0.2, a_1 = 0.3$  and  $c = 1.0, b_1 = 0.4, a_1 = 0.5$ , respectively, and then discard the first 1000 observations. This gives the sample size  $n = 1000$ . I choose the errors  $\epsilon_t$  to have a student's  $t$  distribution with degrees of freedom  $d = 3, 5, 7, 9$ . I truncate likelihood functions at  $\nu = 20$ . I compute coverage probabilities of confidence intervals based on both method I and method II with confidence level 0.90 by taking  $\alpha = 0.99$ . These coverage probabilities are plotted against different sample fraction  $k = 20, 22, \dots, 120$  in Figures 1-4. Our observations from Figures 16 - 19 are as follows:

- 1) These two methods behave similarly, although the normal approximation method may be slightly better. This seems a bit surprising since the data tilting method is better than the normal approximation methods in general. One reason for such an unexpected observation may be the fact that the data tilting method is much more sensitive to the accuracy of estimating innovations than the normal approximation method. This reasoning is confirmed by simulation studies under true innovations, which are not reported here.
- 2) Both methods become accurate when  $d$  becomes large. This is because the tail probability is small for a large  $d$ , i.e., I only need to extrapolate data a little for a large  $d$ .
- 3) In contrast to point estimation, the choice of  $k$  for intervals is more important. This is always difficult, both theoretically and practically. One way to deal with this question is to

have more comparable approaches. Here I propose to choose  $k = 1.5(\log m)^2$ , which is the star points in Figures 16 - 19.

### 4.3.2 The Bootstrap Method

I implement the bootstrap method described in Christoffersen and Goncalves [17]. Table 10 and 11 report the coverage probabilities obtained by the bootstrap method. One can see that the bootstrap method works well for both T3 and T8 distributions with sample size 5,000. However, the bootstrap method is computationally intensive because it requires repetitively solving non-linear optimizations in fitting GARCH models.

**Table 10:** Coverage probabilities by the bootstrap method for 99% one day ahead conditional VaR estimation. Nominal coverage probability level is 90%. The number of simulations is 5,000

<i>SampleSize</i>	1000	5000	8000
<i>T(3)Distribution</i>	0.8414	0.8688	0.8688
<i>T(8)Distribution</i>	0.8966	0.9050	0.9092

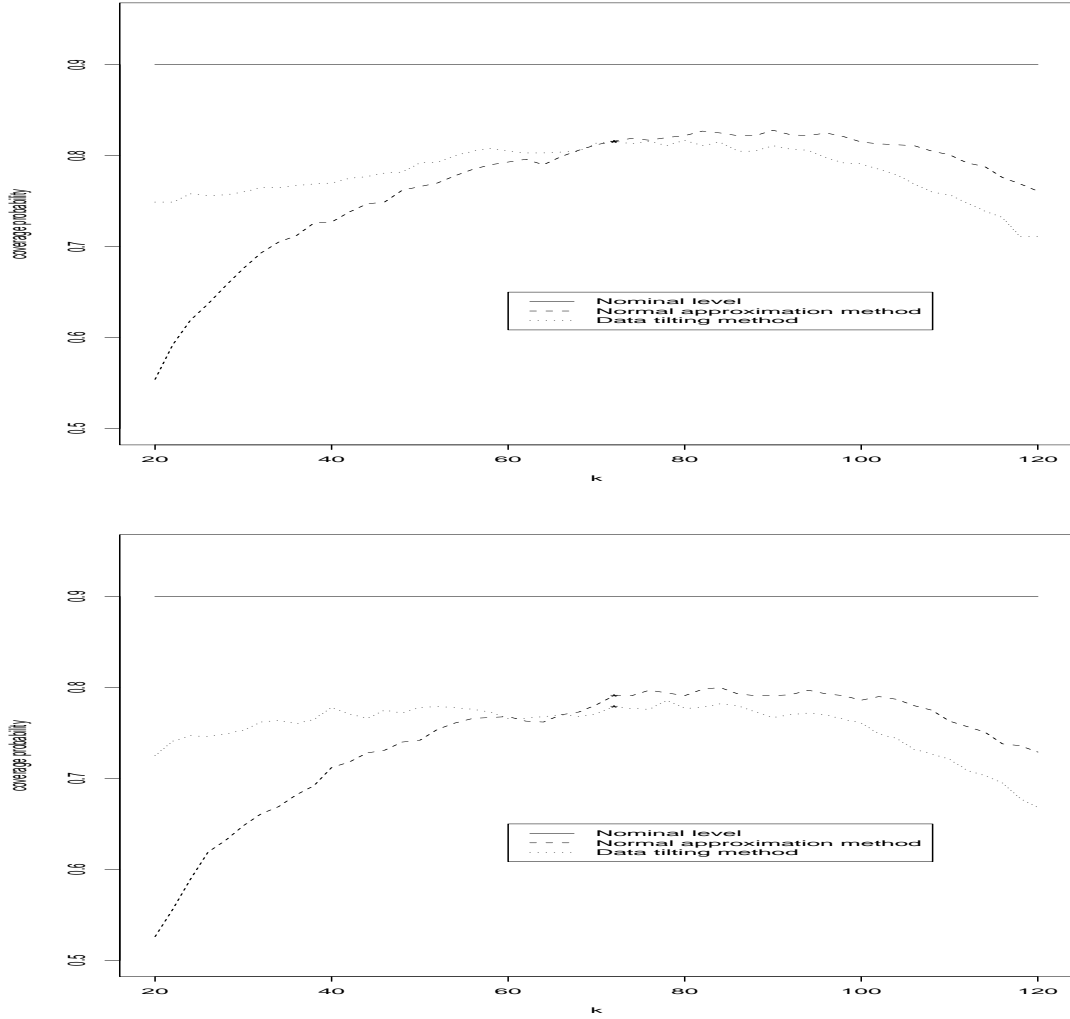
**Table 11:** Coverage probabilities by the bootstrap method for 50% one day ahead conditional VaR estimation. Nominal coverage probability level is 90%. The number of simulations is 5,000

<i>SampleSize</i>	1000	5000	8000
<i>T(3)Distribution</i>	0.8788	0.8836	0.8794
<i>T(8)Distribution</i>	0.9260	0.9086	0.9086

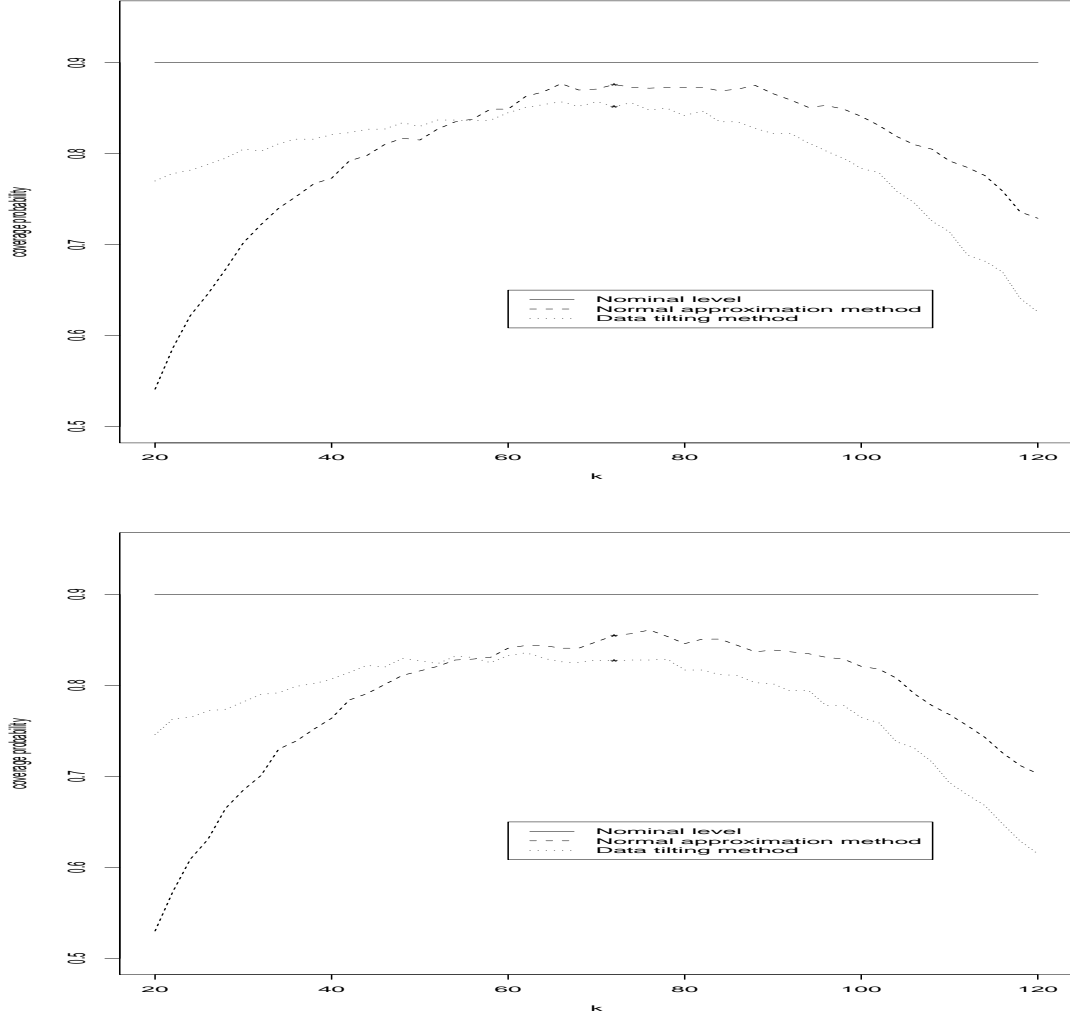
### 4.3.3 Application

I apply our method I to the historical time series data set: the log returns of real time (RT) electricity locational marginal prices (LMPs) in the Pennsylvania-New Jersey-Maryland (PJM) power market from April 1998 to September 2003; see figure 20. Electricity markets are relevantly new markets. Electricity prices in the emerging power markets are much more volatile than prices in most other financial markets due to the almost non-storable nature and the physical production characteristics of electricity.

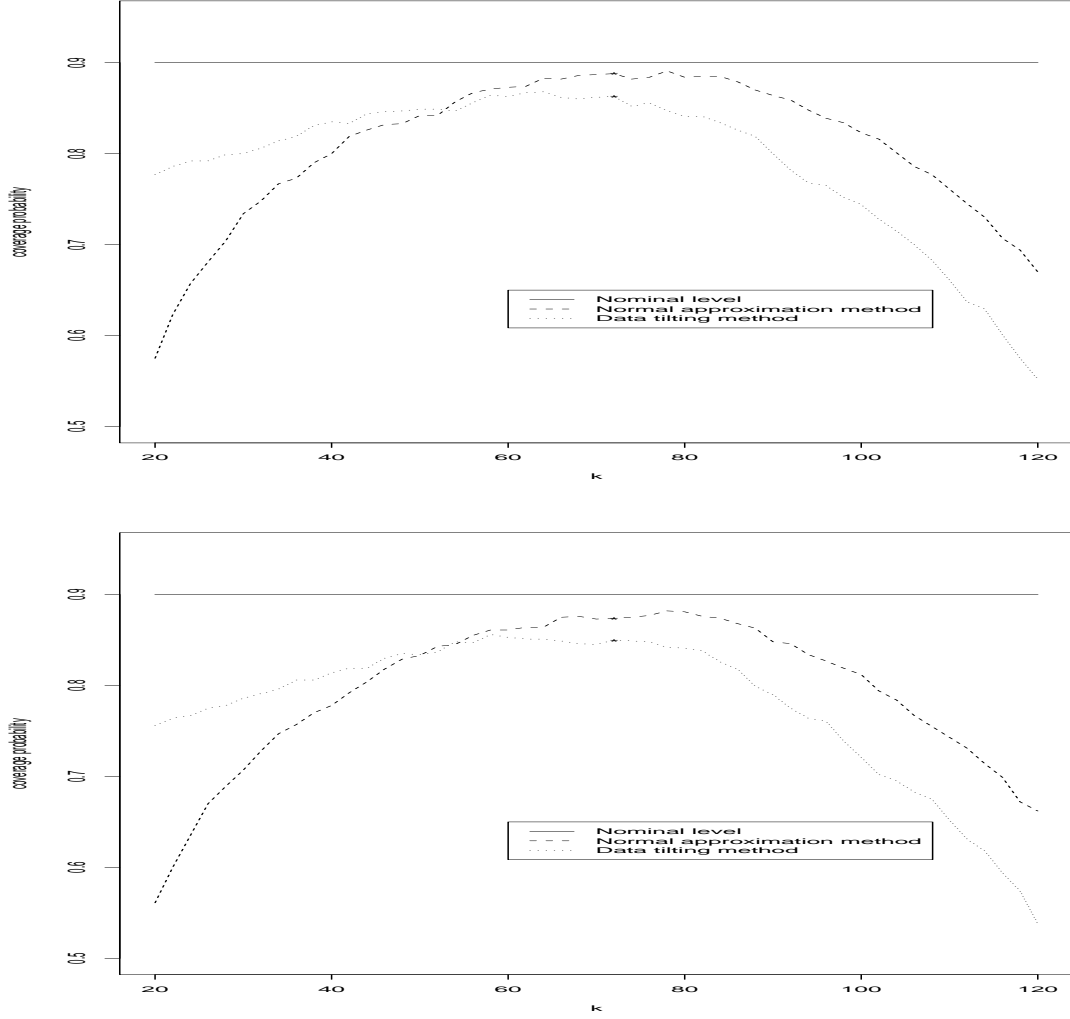
In electricity markets, many market participants, such as utility companies, are especially concerned with the risks of electricity prices rising too high, since they have natural short positions in electricity. In this example, I examine the 99% VaR at the right tails (i.e., the positive-return side) of the conditional distribution of the one-day electricity price return. Like the back test in McNeil and Frey [68], I calculate  $\hat{x}_{\alpha,n}^t$  on day  $t$  in the set  $T = \{n, \dots, N - 1\}$  using a time window of  $n = 500$  days each time where  $N \gg n$ . I run our program for  $\alpha = 0.99$ . For each day  $t \in T$ , I compute the confidence interval based on method I with confidence level 0.90 by taking  $k = 30$ . Furthermore, I pay particular attention to the points that are less than the estimator  $\hat{x}_{\alpha,n}^t$ , but greater than its left endpoint of  $I_\beta^n$  because these points are perceived to have high risks even though they do not exceed our estimated conditional VaR. Thus, I may call the area between the estimator  $\hat{x}_{\alpha,n}^t$  and its left endpoint of  $I_\beta^n$  a “risk-prone” region. Knowing the risk-prone region can be quite valuable in applications such as setting trading limits for traders or evaluating corporate self-insurance exposures since it provides bounds of the conditional VaR estimator at a certain confidence level. I plot log returns of PJM real time LMPs, our estimator and its confidence interval in figure 21 and mark the points in the risk-prone region with squares. In a highly volatile market, conservative market participants may want to employ the interval estimation instead of the point estimator as their VaR estimation.



**Figure 16:** Coverage probabilities for  $t(3)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction  $k = 20, 22, \dots, 120$  for Student  $t$ -distribution with degrees of freedom  $d = 3$ .

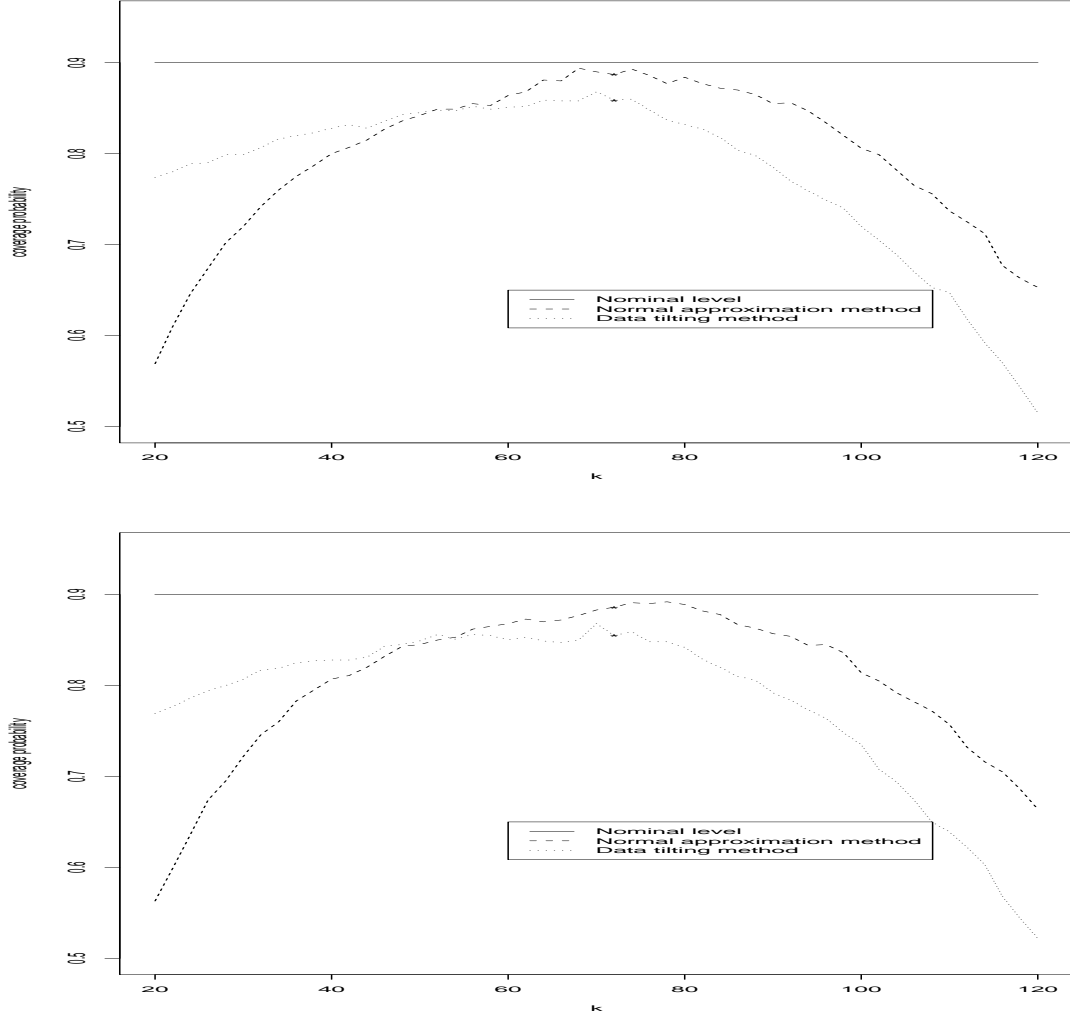


**Figure 17:** Coverage probabilities for  $t(5)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction  $k = 20, 22, \dots, 120$  for Student  $t$ -distribution with degrees of freedom  $d = 5$ .

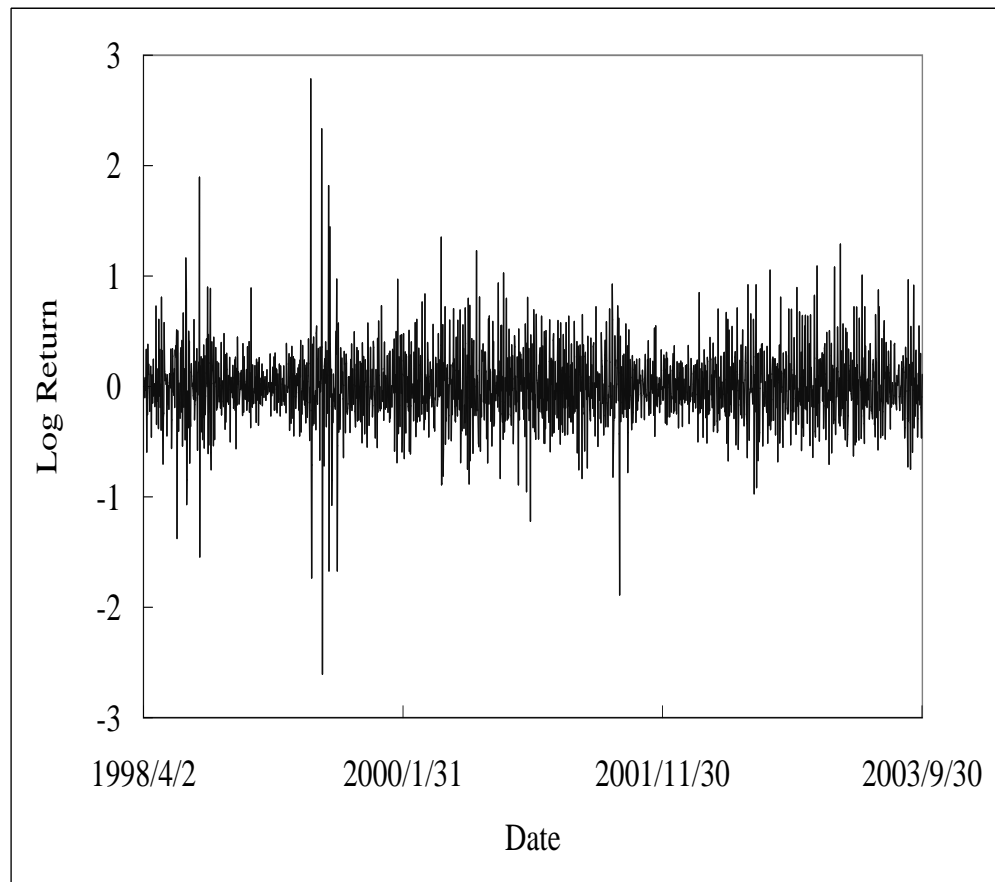


**Figure 18:** Coverage probabilities for  $t(7)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction  $k = 20, 22, \dots, 120$  for Student  $t$ -distribution with degrees of freedom  $d = 7$ .

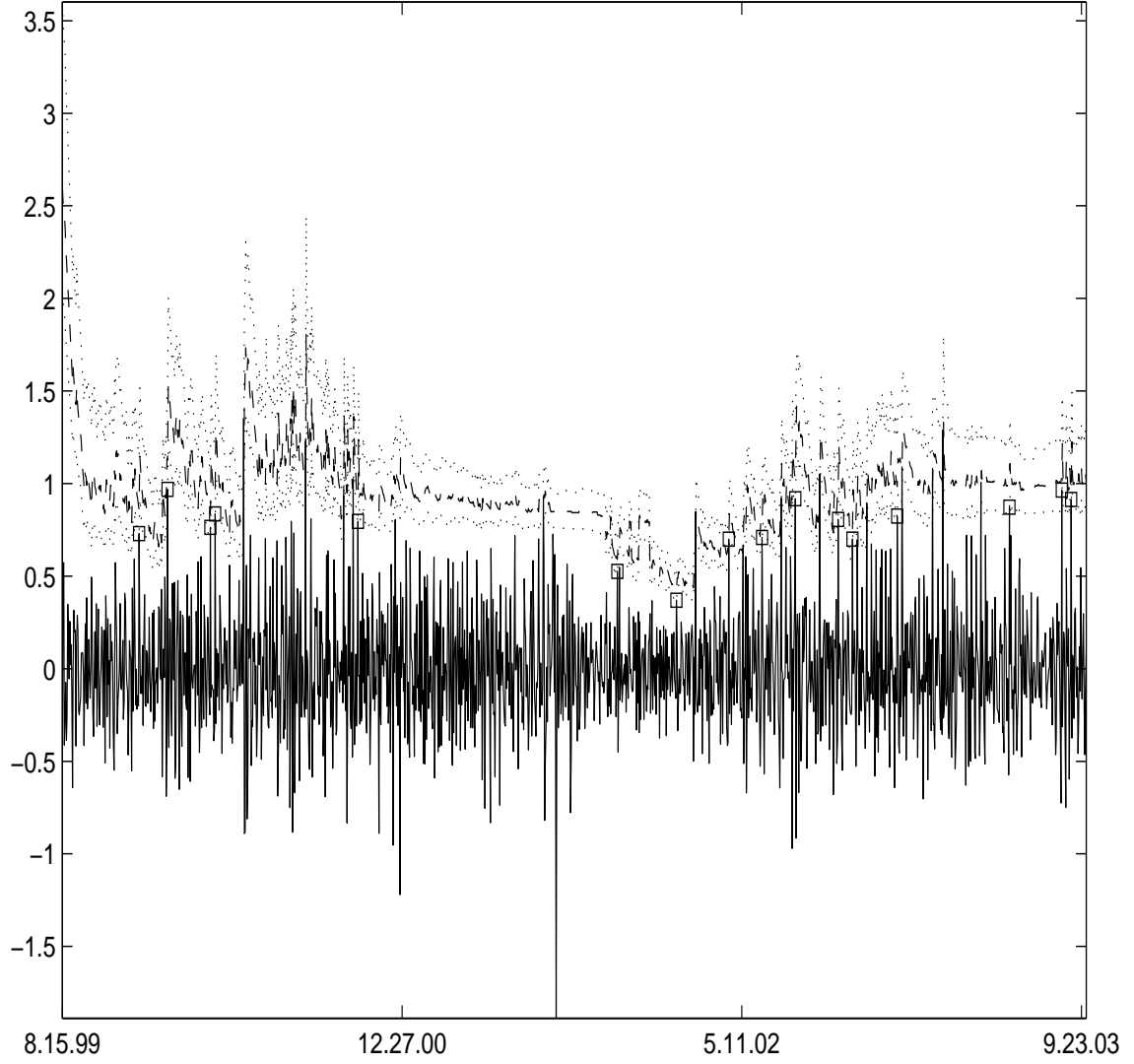




**Figure 19:** Coverage probabilities for  $t(9)$ . The coverage probabilities of confidence intervals based on the normal approximation method and the data tilting method are plotted against different sample fraction  $k = 20, 22, \dots, 120$  for Student  $t$ -distribution with degrees of freedom  $d = 9$ .



**Figure 20:** The log returns of real time electricity locational marginal price in the Pennsylvania - New Jersey - Maryland power market from April 1998 to September 2003.



**Figure 21:** Interval estimation. The estimate  $\hat{x}_{\alpha,n}^t$  (broken line) and the endpoints (dotted lines) of its 90% confidence intervals are plotted against the log return  $X_{t+1}$  (solid line) of PJM real time locational marginal price. I mark those points  $X_{t+1}$  such that  $X_{t+1} < \hat{x}_{\alpha,n}^t$  but greater than the left endpoint of the confidence interval  $I_{\beta}^n$  with squares.

## CHAPTER V

### CONCLUSION AND FUTURE RESEARCH

In this proposal, three issues are addressed: electricity spot prices modeling, electricity supply contract valuation and estimation of the conditional VaR based on a GARCH model with heavy tailed innovations.

Modeling electricity spot prices is a challenging task. To capture the prominent features of electricity prices, I propose a pure jump process with an infinite arrival rate by introducing the mean-reversion feature into the classical variance gamma model. Conditional density function and first four conditional moments are derived, but I encounter some difficulties when estimating the model parameters. Derivatives pricing formulae are derived by transform analysis. As for future research, I feel that it is important to develop an efficient econometric model to perform rigorous parameter estimation, and then apply it to real data.

Electricity tolling agreements, as well as other structured transactions, have played important roles in facilitating risk-sharing and risk-mitigation among independent power producers, utility companies, and unregulated power marketers in the restructured power industry. I propose two simulation-based methods incorporating forward information to value a tolling agreement. Both approaches yield a fairly accurate approximation to the market value of a tolling agreement as gauged by the limitedly available market transaction data. We also examine the effects of different power price assumptions on the tolling contract valuation through the numerical example. In future work, I expect to incorporate more operational and contractual constraints such as minimum downtime and uptime into the valuation model and apply the model to other customized energy supply contracts and physical assets.

I derive the limiting distribution of a high conditional VaR estimator of a family of

GARCH models with heavy-tailed innovations. With the limiting distribution, a traditional normal approximation method is proposed to construct a confidence interval of the conditional VaR estimator. An alternative method for constructing a confidence interval based on the data tilting method is proposed as well. Monte Carlo simulation studies with the GARCH models with Student- $t$  innovations indicate that both methods yield valid confidence intervals for the VaR estimator while the normal approximation has a slightly higher coverage probability. Based on the confidence intervals, one can identify a risk-prone region, which is given by the area between the conditional VaR estimator and the left endpoint of its confidence interval. In practice, one should pay attention to this entire region since it signifies high risk scenarios even though individual points in the region may not exceed the estimated VaR threshold. As a result of the non-parametric setting, the proposed methods are applicable to GARCH models with general innovations including those with asymmetric tails. For instance, they can be applied to asymmetric GARCH models as long as the tail balance assumption (3) holds. Future work is expected to extend these methods to other risk measures such as the expected shortfall probabilities with a broader class of time series models.

# APPENDIX A

## PROOFS IN CHAPTER 4

Throughout this section we shall assume  $p = q = 1$  since other cases can be shown in a similar way. Define

$$\hat{W}_n(u) = k^{-1/2} \sum_{t=\nu}^n \{I(1 - G(\hat{\epsilon}_t) \leq \frac{k}{m}u) - \frac{k}{m}u\}.$$

A lemma first.

**Lemma A.0.1** *As  $n \rightarrow \infty$ ,*

$$\hat{W}_n(u) \xrightarrow{d} B(u) \text{ in } D[0, 1],$$

where  $D[0, 1]$  denotes the space of functions on  $[0, 1]$  which is defined and equipped with the Skorokhod topology (see Billingsley (1968)) and  $\{B(u), u \geq 0\}$  is a standard Brownian motion.

**Proof.** Define

$$\hat{\delta}_{n1} = n\lambda_n^{-1}(\hat{a}_1 - a_1), \quad \hat{\delta}_{n2} = n\lambda_n^{-1}(\hat{b}_1 - b_1), \quad \hat{\delta}_{n3} = n\lambda_n^{-1}(\hat{c} - c),$$

$$s_t(\delta_1, \delta_2, \delta_3) = [\tilde{\sigma}_t(a_1 + n^{-1}\lambda_n\delta_1, b_1 + n^{-1}\lambda_n\delta_2, c + n^{-1}\lambda_n\delta_3) - \sigma_t(a_1, b_1, c)]/\sigma_t(a_1, b_1, c),$$

$$E_{n1}(u, \delta_1, \delta_2, \delta_3) = k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku})(1 + s_t(\delta_1, \delta_2, \delta_3))) - \frac{k}{m}u\},$$

$$\begin{aligned} E_{n2}(u, \delta_1, \delta_2, \delta_3) &= k^{-1/2} \sum_{t=\nu}^n \{I(\epsilon_t \geq U(\frac{m}{ku})(1 + s_2(\delta_1, \delta_2, \delta_3))) \\ &\quad - (1 - G(U(\frac{m}{ku})(1 + s_t(\delta_1, \delta_2, \delta_3)))) + \frac{k}{m}u - I(\epsilon_t \geq U(\frac{m}{ku}))\} \end{aligned}$$

and

$$W_n(u) = k^{-1/2} \sum_{t=\nu}^n \{I(1 - G(\epsilon_t) \leq \frac{k}{m}u) - \frac{k}{m}u\}.$$

Since

$$\hat{W}_n(u) - W_n(u) = E_{n1}(u, \hat{\delta}_{n1}, \hat{\delta}_{n2}, \hat{\delta}_{n3}) + E_{n2}(u, \hat{\delta}_{n1}, \hat{\delta}_{n2}, \hat{\delta}_{n3}),$$

$\hat{\delta}_{n1} = O_p(1)$ ,  $\hat{\delta}_{n2} = O_p(1)$ ,  $\hat{\delta}_{n3} = O_p(1)$  and  $W_n(u) \xrightarrow{D} B(u)$  in  $D[0, 1]$ , to prove this lemma, it is sufficient to show that for any fixed  $\Delta > 0$ ,

$$\sup_{-\Delta \leq \delta_1, \delta_2, \delta_3 \leq \Delta} \sup_{0 \leq u \leq 1} |E_{n1}(u, \delta_1, \delta_2, \delta_3)| = o_p(1) \quad (68)$$

and

$$\sup_{-\Delta \leq \delta_1, \delta_2, \delta_3 \leq \Delta} \sup_{0 \leq u \leq 1} |E_{n2}(u, \delta_1, \delta_2, \delta_3)| = o_p(1). \quad (69)$$

Define

$$\begin{aligned} s_t^*(\Delta) &= s_t(\Delta, \Delta, \Delta) \\ a_{nt}(u, \Delta) &= I(\epsilon_t \geq U(\frac{m}{ku})(1 + s_t^*(\Delta))) \\ &\quad - (1 - G(U(\frac{m}{ku})(1 + s_t^*(\Delta)))) + \frac{k}{m}u - I(\epsilon_t \geq U(\frac{m}{ku})). \end{aligned}$$

Let  $N(n) = [M/A(m/k)]$  for any fixed  $M > 0$ , and  $u_i = i/N(n)$ ,  $i = 0, 1, \dots, N(n)$ . When  $u \in [u_r, u_{r+1}]$ , we have

$$\begin{aligned} & k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta) \\ \leq & k^{-1/2} \sum_{t=\nu}^n a_{nt}(u_{r+1}, \Delta) \\ & + k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku_{r+1}})(1 + s_t^*(\Delta))) - \frac{k}{m}u_{r+1}\} \\ & - k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku_r})(1 + s_t^*(\Delta))) - \frac{k}{m}u_r\} \\ & + 2k^{-1/2} \sum_{t=\nu}^n \{\frac{k}{m}u_{r+1} - \frac{k}{m}u_r\} \\ & + k^{-1/2} \sum_{t=\nu}^n \{I(\epsilon_t \geq U(\frac{m}{ku_{r+1}})) - \frac{k}{m}u_{r+1} + \frac{k}{m}u_r - I(\epsilon_t \geq U(\frac{m}{ku_r}))\} \end{aligned}$$

and

$$\begin{aligned} & k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta) \\ \geq & k^{-1/2} \sum_{t=\nu}^n a_{nt}(u_r, \Delta) \\ & + k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku_r})(1 + s_t^*(\Delta))) - \frac{k}{m}u_r\} \\ & - k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku_{r+1}})(1 + s_t^*(\Delta))) - \frac{k}{m}u_{r+1}\} \\ & + 3k^{-1/2} \sum_{t=\nu}^n \{\frac{k}{m}u_r - \frac{k}{m}u_{r+1}\} \\ & + k^{-1/2} \sum_{t=\nu}^n \{I(\epsilon_t \geq U(\frac{m}{ku_r})) - \frac{k}{m}u_r + \frac{k}{m}u_{r+1} - I(\epsilon_t \geq U(\frac{m}{ku_{r+1}}))\}. \end{aligned}$$

Hence

$$\begin{aligned}
& \sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta)| \\
\leq & \sup_r |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u_r, \Delta)| \\
& + 2 \sup_r |k^{-1/2} \sum_{t=\nu}^n \{1 - G(U(\frac{m}{ku_r})(1 + s_t^*(\Delta))) - \frac{k}{m} u_r\}| \\
& + 3 \sup_r |k^{-1/2} \sum_{t=\nu}^n \{\frac{k}{m} u_{r+1} - \frac{k}{m} u_r\}| \\
& + \sup_r |k^{-1/2} \sum_{t=\nu}^n \{I(\epsilon_t \geq U(\frac{m}{ku_r})) - \frac{k}{m} u_r + \frac{k}{m} u_{r+1} - I(\epsilon_t \geq U(\frac{m}{ku_{r+1}}))\}| \\
= & I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Let  $\mathcal{F}_s = \sigma(\epsilon_t, t \leq s)$ . Then

$$\begin{aligned}
& P(I_1 > \epsilon) \\
\leq & N(n) \sup_r P(|k^{-1/2} \sum_{t=\nu}^n a_{nt}(u_r, \Delta)| > \epsilon) \\
\leq & N(n) k^{-1} \epsilon^{-2} \sup_r E(\sum_{t=\nu}^n a_{nt}(u_r, \Delta))^2 \\
= & N(n) k^{-1} \epsilon^{-2} \sup_r \sum_{t=\nu}^n E\{E(a_{nt}^2(u_r, \Delta) | \mathcal{F}_{t-1})\} \\
\leq & N(n) k^{-1} \epsilon^{-2} \sup_r \sum_{t=\nu}^n E|\frac{k}{m} u_r - (1 - G(U(\frac{m}{ku_r})(1 + s_t^*(\Delta))))| \\
\leq & N(n) k^{-1} \epsilon^{-2} \sup_r \sum_{t=\nu}^n \frac{k}{m} u_r E|1 - (1 + s_t^*(\Delta))^{-\gamma}| \\
& + N(n) k^{-1} \epsilon^{-2} \sup_r \sum_{t=\nu}^n \frac{k}{m} u_r E|\frac{1 - G(U(m/(ku_r))(1 + s_t^*(\Delta)))}{1 - G(U(m/(ku_r)))} - (1 + s_t^*(\Delta))^{-\gamma}| \\
= & II_1 + II_2.
\end{aligned}$$

Set

$$s_{t1}(\delta_1, \delta_2, \delta_3) = (\tilde{\sigma}_t(a_1 + n^{-1} \lambda_n \delta_1, b_1 + n^{-1} \lambda_n \delta_2, c + n^{-1} \lambda_n \delta_3) - \tilde{\sigma}_t(a_1, b_1, c)) / (\sigma_t(a_1, b_1, c)),$$

$$s_{t2} = (\tilde{\sigma}_t(a_1, b_1, c) - \sigma_t(a_1, b_1, c)) / (\sigma_t(a_1, b_1, c)),$$

and let  $D$  denote a generic positive constant. It is easy to check that

$$\left\{ \begin{array}{l} s_t(\delta_1, \delta_2, \delta_3) = s_{t1}(\delta_1, \delta_2, \delta_3) + s_{t2} \\ |s_t^*(\Delta)| \leq D \\ 0 \leq s_{t1}^*(\Delta) \leq D n^{-1} \lambda_n \end{array} \right. \quad (70)$$

and

$$\begin{aligned}
& \sup_{t \geq \nu} E|s_{t2}| \\
\leq & \sup_{t \geq \nu} D E\{b_1 \sum_{j=1}^{\infty} (a_1)^j X_{t-i-j}^2 I(t-1-j < 1)\} \\
\leq & D(a_1)^\nu.
\end{aligned} \quad (71)$$



By (70), (71) and  $\nu/\log n \rightarrow \infty$ , we have

$$II_1 \leq DN(n)(n^{-1}\lambda_n + (a_1)^\nu) \rightarrow 0. \quad (72)$$

Using (67) and Lemma 2 of Draisma et al. (2001) we can show that

$$\sup_r E|((1 - G(U(\frac{m}{ku_r}))(1 + s_t^*(\Delta)))/1 - G(U(\frac{m}{ku_r}))) - (1 + s_t^*(\Delta))^{-\gamma})A^{-1}(\frac{m}{ku_r})| \rightarrow 0. \quad (73)$$

By Potter's inequality (see Geluk and de Haan (1987)), we have

$$\sup_r |u_r A(\frac{m}{ku_r})/A(\frac{m}{k})| \leq D. \quad (74)$$

So, by (73) and (74),

$$II_2 \rightarrow 0. \quad (75)$$

It follows from (72) and (75) that

$$I_1 \xrightarrow{p} 0.$$

Similarly, we can show that

$$\begin{aligned} I_2 &\leq 2 \sup_r k^{-1/2} \sum_{t=\nu}^n \frac{k}{m} u_r | \frac{1 - G(U(m/(ku_r))(1 + s_t^*(\Delta)))}{1 - G(U(m/(ku_r)))} - (1 + s_t^*(\Delta))^{-\gamma} | \\ &\quad + 2 \sup_r k^{-1/2} \sum_{t=\nu}^n \frac{k}{m} u_r |(1 + s_t^*(\Delta))^{-\gamma} - 1| \\ &\xrightarrow{p} 0. \end{aligned} \quad (76)$$

It is easy to show that

$$I_4 \xrightarrow{p} 0, \quad I_3 \rightarrow 0. \quad (77)$$

So

$$\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta)| \xrightarrow{p} 0.$$

Similarly,

$$\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, -\Delta)| \xrightarrow{p} 0.$$

Note that

$$\begin{aligned} &E_{n2}(u, \delta_1, \delta_2, \delta_3) \\ &\leq k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, -\Delta) \\ &\quad + k^{-1/2} \sum_{t=\nu}^n \{(1 - G(U(\frac{m}{ku}))(1 + s_t^*(-\Delta))) - (1 - G(U(\frac{m}{ku}))(1 + s_t^*(\Delta)))\} \end{aligned}$$

and

$$\begin{aligned}
& E_{n2}(u, \delta_1, \delta_2, \delta_3) \\
& \geq k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta) \\
& \quad + k^{-1/2} \sum_{t=\nu}^n \{ (1 - G(U(\frac{m}{ku})(1 + s_t^*(\Delta)))) - (1 - G(U(\frac{m}{ku})(1 + s_t^*(-\Delta)))) \}.
\end{aligned}$$

In a similar way to the proofs of (76) and (77), we can show that

$$\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n G(U(\frac{m}{ku})(1 + s_t^*(\Delta))) - G(U(\frac{m}{ku})(1 + s_t^*(-\Delta)))| \xrightarrow{p} 0.$$

Thus

$$\begin{aligned}
& P(\sup_{-\Delta \leq \delta_1, \delta_2, \delta_3 \leq \Delta} \sup_{0 \leq u \leq 1} |E_{n2}(u, \delta_1, \delta_2, \delta_3)| \geq \epsilon) \\
& \leq P(\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, \Delta)| \geq \epsilon/4) \\
& \quad + P(\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n a_{nt}(u, -\Delta)| \geq \epsilon/4) \\
& \quad + 2P(\sup_{0 \leq u \leq 1} |k^{-1/2} \sum_{t=\nu}^n \{G(U(\frac{m}{ku})(1 + s_t^*(\Delta))) \\
& \quad - G(U(\frac{n}{ku})(1 + s_t^*(-\Delta)))\}| \geq \epsilon/4) \\
& \rightarrow 0,
\end{aligned}$$

i.e., (69) holds. Using the same arguments in the proofs of (76) and (77), it can be seen that (68) holds. Hence the lemma.  $\square$

**Proof of Theorem 4.2.1.** It can be shown by using lemma A.0.1 and the standard arguments in Ferreira, de Haan and Peng (2003).  $\square$

## REFERENCES

- [1] BARZ, G. and JOHNSON, B., “Modeling the prices of commodities that are costly to store: the case of electricity,” *Proceedings of the Chicago Risk Management Conference*, 1998. Chicago, IL.
- [2] BATLLE, C., “A model for electricity generation risk analysis,” *Ph.D dissertation*, 2002. Univ. Pontificia Comillas, Madrid, Spain.
- [3] BATLLE, C. and BARQUÍN, J., “Fuel prices scenario generation based in a multivariate garch model for risk analysis in a wholesale electricity market,” *Proc. 7th Int. Conf. Probabilistic Methods Applied Power Systems*, 2002. Naples, Italy.
- [4] BILLINGSLEY, P., *Convergence of Probability Measures*. New York: Wiley, 1968.
- [5] BLACK, F., “The pricing of commodity contracts,” *Journal of Financial Economics*, vol. 3, pp. 167–179, 1976.
- [6] BLACK, F. and SCHOLES, M. S., “The pricing of options and corporate liabilities,” *Journal of Political Economy*, vol. 81, pp. 637–654, 1973.
- [7] BOLLERSLEV, T., “Generalized autoregressive conditional heteroscedasticity,” *Journal of Econometrics*, vol. 31, pp. 307–327, 1986.
- [8] BOLLERSLEV, T., CHOU, R., and KRONER, K., “Arch modelling in finance. a review of the theory and empirical evidence,” *Journal of Econometrics*, vol. 52, pp. 5–59, 1992.
- [9] BOYLE, P. P., “A lattice framework for option pricing with two state variables,” *Journal of Financial & Quantitative Analysis*, no. 23, pp. 1–12, 1988.
- [10] BREEDEN, D. and LITZENBERGER, R., “Prices of state contingent claims implicit in option prices,” *Journal of Business*, no. 51, pp. 621–651, 1978.
- [11] BUNN, D. W., *Modeling Prices in Competitive Electricity Markets*. Chichester: John Wiley and Sons, 2004.
- [12] CARR, P. and MADAN, D. B., “Optimal positioning in derivative securities,” *Quantitative Finance*, 2001.
- [13] CARR, P., “An alternative approach for valuing and hedging a corridor call contract,” *Working Paper*, 2003.
- [14] CARR, P., LIPTON, A., and MADAN, D., “An alternative approach for valuing continuous cash flows,” *Working Paper*, 2001.
- [15] CARRIERE, J. F., “Valuation of the early-exercise price for derivative securities using simulations and nonparametric regression,” *Insurance: Mathematics and Economics*, vol. 19, pp. 19–30, 1996.

- [16] CHENG, A. R., GALLANT, A. R., JI, C. S., and LEE, B., “A central limit theorem for computation of option prices for stochastic volatility models,” *Working Paper*, 2004.
- [17] CHRISTOFFERSEN, P. and GONÇALVES, S., “Estimation risk in financial risk management,” *Journal of Risk*, vol. 7, no. 3, 2005.
- [18] CLEWLOW, L. and STRICKLAND, C., *Energy Derivatives - Pricing and Risk Management*. Lacima, 2000.
- [19] COX, J., ROSS, S., and RUBINSTEIN, M., “Option pricing: A simplified approach,” *Journal of Financial Economics*, vol. 7, no. 3, pp. 229–263, 1979.
- [20] CULLITY, B. D., *Introduction to Magnetic Materials*. Reading, MA: Addison-Wesley, 1972.
- [21] DANIELSSON, J. and VRIES, C. D., “Value-at-risk and extreme returns,” *FMG-Discussion Paper No.273*, 1997. Financial Markets Group, London School of Economics.
- [22] DEANS, S. R., *The Radon Transform and Some of Its Applications*. New York: Wiley, 1983.
- [23] DENG, S. J., “Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes,” *POWER Working Paper*, 1999. University of California Energy Institute.
- [24] DENG, S. J., “Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes,” *Working paper*, 2000. Georgia Institute of Technology.
- [25] DENG, S. J., JOHNSON, B., and SOGOMONIAN, A., “Exotic electricity options and the valuation of electricity generation and transmission assets,” *Decision Support Systems*, vol. 30, no. 3, pp. 383–392, 2001.
- [26] DENG, S. J. and LEE, S., “Pricing american-style options with adaptive simulation,” *Working Paper*, 2005. Georgia Institute of Technology.
- [27] DENG, S. J. and OREN, S. S., “Option-based valuation of a power plant incorporating physical constraints,” *Working Paper*, 1998. University of California at Berkeley.
- [28] DENG, S. J. and OREN, S. S., “Incorporating operational characteristics and start-up costs in option-based valuation of power generation capacity,” *Probability in the Engineering and Informational Sciences*, vol. 17, no. 2, pp. 155–181, 2003.
- [29] DICICCIO, T. J., HALL, P., and ROMANO, J. P., “Bartlett adjustment for empirical likelihood,” *Ann. Statist.*, vol. 19, pp. 1053–1061, 1991.
- [30] DIXIT, A. K. and PINDYCK, R. S., *Investment under Uncertainty*. New Jersey: Princeton University Press, 1994.
- [31] DOWD, K., *Measuring Market Risk*. Chichester: Wiley, 2002.

- [32] DRAISMA, G., HAAN, L. D., PENG, L., and PEREIRA, T. T., “A bootstrap-based method to achieve optimality in estimating the extreme-value index,” *Extremes*, vol. 2, pp. 367–404, 1999.
- [33] DUFFIE, D. and GRAY, S., *Volatility in Energy Prices. Manage Energy Price Risk, 2nd edition*. London: Risk Publications, 1998.
- [34] DUFFIE, D. and PAN, J., “An overview of value at risk,” *Journal of Derivative*, vol. spring, pp. 7–48, 1997.
- [35] DUFFIE, D., PAN, J., and SINGLETON, K., “Transform analysis and asset pricing for affine jump-diffusions,” *Econometrica*, vol. 68, no. 6, pp. 1343–1376, 2000.
- [36] ENGLE, R. F., “Autoregressive conditional heteroscedasticity with estimates of the variance of u.k. inflation,” *Econometrica*, vol. 50, pp. 987–1008, 1982.
- [37] ERAKER, B., “Do stock prices and volatility jump? reconciling evidence from spot and option prices,” *Journal of Finance*, vol. 54, no. 3, pp. 1376–1404, 2004.
- [38] FERREIRA, A., HAAN, L. D., and PENG, L., “On optimising the estimation of high quantiles of a probability distribution,” *Statistics*, vol. 37, pp. 403–434, 2003.
- [39] GELUK, J. and HAAN, L. D., *Regular Variation, Extensions and Tauberian Theorems*. Amsterdam: CWI Tract 40, 1987.
- [40] GILBARG, D. and TRUDINGER, N. S., *Elliptic Partial Differential Equations of Second Order*. Berlin, Germany: Springer-Verlag, 1998.
- [41] GILBERT, R. and HENLY, J., “The value of rate reform in a competitive electric power market,” *Regulatory Choices: A Perspective on Developments in Energy Policy*, 1991. Berkeley: University of California Press.
- [42] GONZÁLEZ, A., ROQUE, A. M. S., and GARCÍA-GONZÁLEZ, J., “Modeling and forecasting electricity prices with input/output hidden markov models,” *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 13–24, 2005.
- [43] GOTO, M. and KAROLYI, G. A., “Understanding electricity price volatility within and across markets,” *Working paper*, 2003. Ohio State University.
- [44] GREEN, R. C. and JARROW, R. A., “Spanning and completeness in markets with contingent claims,” *Journal of Economic Theory*, vol. 41, pp. 202–210, 1987.
- [45] HALL, P. and SCALA, B. L., “Methodology and algorithms of empirical likelihood,” *International Statistical Review*, vol. 58, no. 2, pp. 109–127, 1990.
- [46] HALL, P. and YAO, Q., “Data tilting for time series,” *Journal of the Royal Statistical Society*, vol. Series B 65, pp. 425–442, 2003.
- [47] HALL, P. and YAO, Q., “Inference in arch and garch models with heavy-tailed errors,” *Econometrica*, vol. 71, pp. 285–317, 2003.
- [48] HAMILTON, J. D., *Time Series Analysis*. Princeton, New Jersey: Princeton University Press, 1994.

- [49] HANSEN, L. P., “Large sample properties of generalized method of moments estimators,” *Econometrica*, vol. 50, pp. 1029–1054, 1982.
- [50] HARRISON, J. M. and KREPS, D., “Martingales and arbitrage in multiperiod securities markets,” *Journal of Economic Theory*, vol. 20, pp. 381–408, 1979.
- [51] HARRISON, J. M. and PLISKA, S., “Martingales and stochastic integrals in the theory of continuous trading,” *Stochastic Processes and Their Applications*, vol. 11, pp. 215–260, 1981.
- [52] HE, H., “Convergence from discrete- to continuous-time contingent claims prices,” *The Review of Financial Studies*, vol. 3, no. 4, pp. 523–546, 1990.
- [53] HIGHAM, D. J., “An algorithmic introduction to numerical simulation of stochastic differential equations,” *SIAM Review*, vol. 43, no. 3, pp. 525–546, 2001.
- [54] HILL, B., “A simple approach to inference about the tail of a distribution,” *Annals of Statistics*, vol. 3, pp. 1163–1174, 1975.
- [55] HORNIK, K., STINCHCOMBE, M., and WHITE, H., “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [56] JAILLET, P., RONN, E., and TOMPAIDIS, S., “Valuation commodity based swing options,” *Working Paper*, 2001. University of Texas, Austin.
- [57] JORION, P., *Value-at-Risk: The New Benchmark for Managing Financial Risk*. New York: McGraw-Hill, 2000.
- [58] KAMINSKI, V., “The challenge of pricing and risk managing electricity derivatives,” *The US Power Market*, pp. 149–171, 1997. Risk Publications.
- [59] KARATZAS, I. and SHREVE, S., *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag, 1988.
- [60] KEPPO, J., “Pricing of electricity swing options,” *Working Paper*, 2004. University of Michigan.
- [61] KNITTEL, C. R. and ROBERTS, M. R., “An empirical examination of deregulelectricity prices,” *working paper*, 2001. the Program on Workable Energy Regulation (POWER), University of California Energy Institute, Berkeley, California.
- [62] LIPTON, A., *Mathematical Methods For Foreign Exchange: A Financial Engineer’s Approach*. Singapore: World Scientific, 2001.
- [63] LONGSTAFF, F. A. and SCHWARTZ, E. S., “Valuing american options by simulation: A simple least-squares approach,” *The Review of Financial Studies*, vol. 14, no. 1, pp. 113–147, 2001.
- [64] LONGSTAFF, F. A. and WANG, A. W., “Electricity forward prices: a high-frequency empirical analysis,” *Journal of Finance*, vol. 59, no. 4, pp. 1877–1900, 2004.
- [65] LUCIA, J. J. and SCHWARTZ, E. S., “Electricity prices and power derivatives: evidence from the nordic power exchange,” *Review of Derivatives Research*, vol. 5, no. 1, pp. 5–50, 2002.

- [66] MADAN, D. B., CARR, P. P., and CHANG, E. C., "The variance gamma process and option pricing," *European Finance Review*, vol. 2, pp. 79–105, 1998.
- [67] MARGRABE, W., "The value of an option to exchange one asset for another," *Journal of Finance*, vol. 33, no. 1, pp. 177–186, 1978.
- [68] MCNEIL, A. and FREY, R., "Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach," *Journal of Empirical Finance*, vol. 7, pp. 271–300, 2000.
- [69] MEINSHAUSEN, N. and HAMBLY, B. M., "Monte carlo methods for the valuation of multiple-exercise options," *Mathematical Finance*, vol. 14, no. 4, pp. 557–583, 2004.
- [70] MERTON, R. C., "Theory of rational option pricing," *Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.
- [71] MITTNIK, S. and RACHEV, S. T., *Stable Paretian Models in Finance*. New York: Wiley, 2000.
- [72] MOUNT, T. and ETHIER, R., "Estimating the volatility of spot prices in restructured electricity markets and the implications for option values," *P SERC Working paper*, 1998. Cornell University.
- [73] NACHMAN, D. C., "Spanning and completeness with options," *The Review of Financial Studies*, vol. 1, no. 3, pp. 311–328, 1988.
- [74] NEAL, R. M., "Probabilistic inference using markov chain monte carlo methods," *Technical Report*, 1993. Department of Computer Science, University of Toronto.
- [75] NEWEY, W. K. and WEST, K. D., "A simple positive semi-definite, heteroskedasticity and autocorrelation consistent matrix," *Econometrica*, vol. 55, pp. 703–708, 1987.
- [76] OWEN, A., "Empirical likelihood ratio confidence intervals for single functional," *Biometrika*, vol. 75, pp. 237–249, 1988.
- [77] OWEN, A., "Empirical likelihood region," *Biometrika*, vol. 75, pp. 237–249, 1988.
- [78] OWEN, A., *Empirical Likelihood*. New York: Chapman and Hall, 2001.
- [79] PENG, L. and QI, Y., "Confidence intervals for high quantiles of a heavy tailed distribution," *Technical Report*, 2003. Georgia Institute of Technology, Atlanta.
- [80] PENG, L. and YAO, Q., "Least absolute deviations estimator for arch and garch models," *Biometrika*, vol. 90, pp. 967–975, 2003.
- [81] ROSS, S., "Options and efficiency," *Quarterly Journal of Economics*, vol. 90, pp. 75–89, 1976.
- [82] SCHWARTZ, E. and SMITH, J. E., "Short-term variations and long-term dynamics in commodity prices," *Management Science*, vol. 46, no. 7, pp. 893–911, 2000.
- [83] SCHWARTZ, E. S., "The stochastic behavior of commodity prices: Implications for valuation and hedging," *Journal of Finance*, vol. 52, pp. 923–973, 1997.

- [84] SCHWARTZ, E. S., ed., *Real Options and Investment Under Uncertainty*. Cambridge, Mass.: MIT Press, 2001.
- [85] SCHWEPPE, F. C., CARAMANIS, M. C., TABORS, R. D., and BOHN, R. E., *Spot Pricing of Electricity*. Boston: Kluwer Academic Publishers, 1988.
- [86] SHIMKO, D., *VaR for Corporates. VaR: Understanding and Applying Value-at-Risk*. London: Risk Publications, 1997.
- [87] TAYLOR, S. J., *Modelling Financial Time Series*. New York: Wiley, 1986.
- [88] THOMPSON, A., "Valuation of path-dependent contingent claims with multiple exercise decisions over time: the case of take-or-pay," *Journal of Quantitative and Computative Analysis*, vol. 30, no. 2, pp. 271–293, 1995.
- [89] TRIGEORGIS, L., ed., *Real Options in Capital Investment: Models, Strategies, and Applications*. Westport, CT, and London: Praeger, 1995.
- [90] TSENG, C.-L. and BARZ, G., "Short-term generation asset valuation: a real options approach," *Operations Research*, vol. 50, no. 2, pp. 297–310, 2002.
- [91] TSITSIKLIS, J. N. and VAN ROY, B., "Regression methods for pricing complex american-style options," *IEEE Transactions on Neural Networks*, forthcoming, 2001.
- [92] VENTOSA, M., BAÍLLO, A., RAMOS, A., and RIVIER, M., "Electricity market modeling trends," *Energy Policy*, 2004. in press.
- [93] WOO, C.-K., OLSON, A., and ORANS, R., "Benchmarking the price reasonableness of an electricity tolling agreement," *Electricity Journal*, vol. 17, no. 5, pp. 65–75, 2004.
- [94] WOOD, A. and WOLLENBERG, B., *Power Generation, Operation and Control*. New York: John Wiley & Sons, 1984.
- [95] WORTHINGTON, A. C., KAY-SPRATLEY, A., and HIGGS, H., "Transmission of prices and price volatility in australian electricity spot markets: a multivariate garch analysis," *Energy Economics*, vol. 27, no. 2, pp. 337–350, 2003.
- [96] ZHOU, H., "Itô conditional moment generator and the estimation of short rate processes," *Working Paper*, 2003. Federal Reserve Board.